

THÈSE DE DOCTORAT DE L'UNIVERSITÉ PARIS 6

Ecole doctorale EDITE

Spécialité

Acoustique, Traitement du signal, Informatique Appliqués à la Musique

présentée par

Nicolas Hainiandry RASAMIMANANA

Nicolas.Rasamimanana@ircam.fr

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE PARIS 6

Geste instrumental
du violoniste en situation de jeu :
analyse et modélisation.

effectuée sous la direction de : René CAUSSÉ

soutenue le 10/03/2008

membres du jury :

René CAUSSÉ	Directeur de thèse
Marc LEMAN	Rapporteur
Marcelo WANDERLEY	Rapporteur
Frédéric BEVILACQUA	Examineur
Benoît FABRE	Examineur
Sylvie GIBET	Examineur

IRCAM-Centre Pompidou
CNRS, UMR STMS
1, Place Igor Stravinsky
75004 PARIS - FRANCE.

Résumé

Les techniques de captation de gestes offrent aux interprètes la possibilité de contrôler des processus de synthèse sonore par leurs mouvements. Les réflexions sur les manières d'associer "geste" et "son" amènent alors à s'interroger sur les liens essentiels unissant un musicien et son instrument. Dans ce contexte, l'étude d'instruments acoustiques et des gestes employés par les instrumentistes pour contrôler le son émis peut apporter des éléments de réponse fondamentaux.

Cette thèse étudie ce contrôle dans le cas des cordes frottées, en se fixant dans un cadre défini par les contraintes acoustiques de l'instrument et biomécaniques de l'instrumentiste. Elle s'articule autour de quatre articles. Les articles I et III proposent une exploration des stratégies utilisées par les instrumentistes dans différents contextes impliquant différents modes de jeux, tempos et nuances. L'article II se concentre sur les transitions entre coups d'archet, produisant différentes qualités d'articulations entre notes, et souligne leur importance dans la maîtrise du contrôle de l'archet. L'article IV propose des modèles du mouvement inspirés de travaux en biomécanique, et permettant notamment de mettre en avant l'influence du contexte musical.

Les applications à la création artistique et à la pédagogie musicale sont également présentées.

Mots clés : analyse du geste, mouvement, instruments à cordes frottées, modélisation du geste, performance musicale.

Abstract

Gesture sensing technologies allow musicians to control sound synthesis processes through their body movements. Reflections on how to map "gestures" and "sound" therefore lead to question the essential links between a musician and his/her instrument. In this context, the study of acoustic instruments and of performers gestures used to control acoustic sounds can bring fundamental understanding for electronic music performance.

This thesis studies such a control in the case of bowed strings, using a framework defined by the instrument's acoustic constraints and the player's biomechanical constraints. Results are presented in four articles. Papers I and III propose to explore various strategies used by players in different contexts involving several bowing techniques, tempos and nuances. Paper II focuses on transitions between bow strokes, producing different qualities of note articulations, and underlines their importance in bow mastery. Paper IV presents movement models inspired from biomechanical works and notably putting forth the influence of musical context.

Applications of the results to music creation and music pedagogy are also presented.

Keywords : gesture analysis, movement, bowed string instruments, gesture modelling, music performance.

Remerciements

Ces trois années de thèse ont été très riches pour moi, tant sur le plan intellectuel que personnel. Je le dois à un ensemble de circonstances et de personnes qui se sont trouvées en phase avec mes attentes et qui ont su par ailleurs en provoquer de nouvelles. Il faudrait pouvoir remercier cet ensemble en tant que tel, comme ça, d'un coup. Pourtant bien que profondément sincère, cela laisserait un arrière goût de remerciements bâclés. Je me lance donc dans une énumération, certes plus traditionnelle mais toujours aussi périlleuse.

Je tiens tout d'abord à remercier les membres de mon jury et en particulier les rapporteurs Marc Leman et Marcelo Wanderley, pour avoir pris le temps de lire mon manuscrit et d'évaluer mon travail de thèse. Je remercie également René Caussé, qui a accepté de diriger cette thèse et qui a aussi accepté de laisser l'encadrement à Frédéric Bevilacqua. Un grand merci à Norbert Schnell qui m'a chaleureusement accueilli au sein de l'équipe Applications Temps Réel, aujourd'hui rebaptisée Interactions Musicales Temps Réel, équipe dont j'ai l'honneur d'être le premier doctorant. Merci également au projet européen i-Maestro et au projet ANR Consonnes pour leurs soutiens, et merci au Centre de Coopération InterUniversitaire France-Québec pour m'avoir aidé pendant mon séjour à Montréal.

Plusieurs personnes ont eu une influence significative et positive sur mon travail de thèse. En premier lieu, je tiens à dire un grand, grand merci à Frédéric Bevilacqua pour m'avoir guidé pendant cette thèse, pour avoir su me conseiller tout en me laissant une grande liberté d'exploration. Merci d'avoir créé les conditions de recherche propices. Je tiens aussi à remercier Marcelo Wanderley qui a toujours montré de l'enthousiasme pour mon travail et dont les conseils m'ont toujours été très utiles. Je remercie également mes deux collègues de recherche Erwin Schoonderwaldt et Matthias Demoucron pour nos collaborations passées et à venir. Merci aussi à Delphine Bernardin pour ses conseils sur les aspects contrôle moteur. Merci à Florian Kaiser pour son aide précieuse sur l'analyse des changements d'archet au violon. Je tiens enfin à remercier tous les membres de l'équipe IMTR pour leurs grandes compétences. Merci en particulier à Arshia Cont pour son œil extérieur.

Ce que je retiens surtout de mon séjour à l'IRCAM, ce sont les personnes aux parcours diverses que cet institut rassemble. Cette richesse de personnes m'a permis de trouver inspirations, idées, critiques, motivations, conseils en provenance de différents horizons. Pour cela, je tiens à remercier Philippe Manoury, Hector Parra, Christophe Desjardins, Serge LeMouton, Hae-Sun Kang, Emmanuel Fléty, Franck Bedrossian, Eric Daubresse, Jérôme Combier, Nicolas Leroy, Anne Mercier, Sebastian Rivas, David Poissonnier, Jeanne-

Marie Conquer. Je remercie tout particulièrement Florence Baschet pour les moments forts de *BogenLied*, pour sa compréhension profonde des enjeux de la recherche scientifique et pour nos collaborations très riches. Un grand merci à tous les ircamiens pour leur passion communicative.

Comme je l'ai dit auparavant, ces trois années ont été non seulement riches intellectuellement mais aussi riches d'un point de vue personnel. Pour cela, je tiens particulièrement à remercier trois personnes avec qui j'ai pu partager plus qu'un bout de route et qui malgré tous mes efforts ont réussi à devenir plus que de simples collègues ;-). Merci à Matthias Demoucron pour toutes nos discussions, pour les plus sensées et aussi pour les plus absurdes. Merci de me ressembler, un peu. Merci à Damien Tardieu pour sa grande qualité d'écoute, de tout contenu, en toutes circonstances. Merci de détester les pigeons. Un immense merci à Julien Bloit pour ses qualités humaines rares, pour avoir réussi à m'humaniser un peu. Merci de m'avoir fait découvrir les sphynx. Un grand merci à vous trois pour tout, vraiment. Merci également à Alice Daquet pour avoir abondé dans mes folies. Je tiens aussi à remercier particulièrement Denis Matignon qui m'a écouté et éclairé à un moment charnière de ma vie.

Ces années de thèse ont été l'occasion pour moi de renouer avec une pratique musicale plus soutenue. Par rapport à ma recherche, cela m'a permis de jongler avec plus de facilités entre casquette de chercheur et casquette d'instrumentiste. Un très grand merci à Ruxandra Sirli pour ses bons conseils, pour sa bonne humeur et pour m'avoir emmené plus loin dans ma pratique musicale. Un merci infini à mon luthier Patrick Arguence pour sa grande gentillesse, pour son humour et pour avoir toujours remédié efficacement à mes petits soucis psycho-patho-acoustiques. Merci aussi pour les heures passées dans l'atelier à discuter de tout sauf de mon violon ! Merci à Fabrice Guédy de m'avoir accueilli en temps que professeur de violon à l'atelier des Feuillantines, merci pour cette confiance. Merci également à tous mes partenaires de musique de chambre. Merci à Benjamin Leclaire pour les sonates piano - violon, à Béatrice Puysségur pour les duos de violons, à Jean-Michel Levy-Bruhl pour les trios piano - violon - violoncelle. Merci également à Mathieu pour les quelques moments musicaux autour de Beethoven et merci à Matthias pour les rares fois où nous avons réussi à jouer ensemble ;-).

Je tiens également à remercier ma famille qui m'a soutenu moralement tout au long de ces années. Merci, vraiment. Merci aussi à tous mes amis pour les bons moments qui changent les idées. Enfin, je ne saurais assez remercier David et Dora, pour les petits instants, pour les grands, pour la joie de vivre et le bonheur qu'ils m'apportent jour après jour. Ca m'est vital. Merci.

*Le fou copie l'artiste, et l'artiste ressemble au fou.
(André Malraux, L'Espoir).*

Je dédie cette thèse à ma mère, à son courage.

Table des matières

I	Introduction	11
1	Introduction	13
1.1	Ce que l'on entend par <i>geste</i> : définitions	15
1.1.1	Nommer les gestes musicaux : état de l'art	15
1.1.2	Gestes de <i>production</i> et contrôle de l'archet	16
1.2	De l'intérêt de modéliser le geste	17
1.2.1	A la recherche de la corporalité perdue	17
1.2.2	La norme MIDI : séparation historique du geste et du son	20
1.2.3	Le cas des instruments à son entretenu	21
1.2.4	Musique mixte et interaction musicale à l'IRCAM	22
1.3	Problématique de la thèse	24
2	Contribution	27
2.1	Mesures des gestes du violoniste en situation de jeu : métho- dologie	29
2.1.1	Principaux paramètres physiques des cordes frottées . .	29
2.1.2	Capture des paramètres de l'archet	31
2.1.3	Visualisations des données	44
2.1.4	Des contextes musicaux et non-musicaux : protocole . .	46
2.2	Article I	48
2.3	Article II	51
2.4	Article III	56
2.5	Article IV	59

3	Discussion générale et applications	65
3.1	Discussion générale	65
3.2	Applications	77
4	Directions futures et conclusion	83
4.1	Suivi de performance	83
4.2	Contrôle de la synthèse sonore	84
4.3	Interactions entre interprètes	85
4.4	Conclusion	86
	Bibliographie	87
II	Articles	105
III	Annexes	171

Première partie

Introduction

Chapitre 1

Introduction

Sur la nouvelle scène de la salle Pleyel, David Grimal entre, se place face au public, pose son violon sur son épaule, se concentre. Son archet s'élève et attaque la corde sol marquant ainsi le début de *Tzigane, Rhapsodie de concert* de Maurice Ravel. Pendant une dizaine de minutes, l'auditoire va pouvoir écouter la performance du violoniste et également la voir : les mouvements du corps de l'interprète accompagnant la musique, les mouvements de son archet contrôlant la vibration des cordes, de sa main gauche donnant la hauteur des notes. Pendant l'interprétation du musicien, le public ne sera pas moins sensible à la dimension visuelle qu'à la dimension auditive. Contrairement à l'écoute d'un enregistrement, l'auditeur pourra voir quelle est l'origine des sons qu'il entend et pourra faire le lien entre l'énergie donnée par l'interprète et les sons émis par l'instrument. Ainsi, les gestes effectués par David Grimal lors de ce récital participent entièrement à la compréhension du discours musical et à l'engagement du public dans la performance. Conscients de ce phénomène, les acteurs en informatique musicale, chercheurs et musiciens, se sont interrogés sur les possibilités d'engager les mouvements du corps d'un "interprète" pour contrôler des sons résultant généralement d'un algorithme informatique, cherchant ainsi à recréer artificiellement les liens essentiels unissant un musicien et son instrument. Ainsi, ces dernières années, plusieurs problématiques autour de cette possibilité ont émergé en informatique musicale comme en attestent par exemple la création de la conférence

internationale *New Interfaces for Musical Expression* (NIME) dont la sixième édition s’est déroulée à l’IRCAM en 2006, ou la part grandissante accordée au contrôle gestuel dans les plus anciennes conférences telles que *International Computer Music Conference* (ICMC), ou encore l’édition de premiers ouvrages de référence comme (Wanderley and Battier, 2000).

Les enjeux sont multiples. Les avancées conséquentes qu’ont connues les techniques de synthèse sonore depuis leur début (synthèse additive, par modèles spectraux, synthèse granulaire, synthèse concaténative, synthèse par modèle physique) sont aujourd’hui confrontées au problème du contrôle de leurs paramètres d’entrées (Verfaillie et al., 2006b), paramètres parfois très nombreux (modèles spectraux) et/ou dont l’évolution temporelle peut être complexe (modèles physiques). Parallèlement, les réalisateurs de nouvelles interfaces permettant la capture de *gestes* d’un opérateur humain sont aujourd’hui moins confrontés au problème d’implémentation des capteurs électroniques qu’au problème d’associer leurs données captées à des sons ou des paramètres sonores pour obtenir un résultat convainquant (Hunt et al., 2003). Les réflexions menées autour de ces deux problèmes convergent cependant vers une même question : est-il possible de trouver un moyen d’interagir avec des machines de manière à ce que l’expression artistique ne fasse qu’un avec les technologies contemporaines ? (Leman, 2007, Introduction). Cette question vient fixer le contexte général dans lequel s’inscrit la recherche menée pendant cette thèse.

Dans ce cadre de recherche, l’étude des instruments acoustiques et de la manière dont les musiciens les manipulent peut apporter des éléments de réponses fondamentaux (Leman, 2007, page 138). En se plaçant dans cette ligne, les travaux rapportés dans ce manuscrit traitent de l’étude des instruments à cordes frottées et des *gestes* employés par les musiciens pour contrôler le son émis en situation de jeu.

1.1 Ce que l'on entend par *geste* : définitions

Il existe plusieurs usages du terme *geste* dans la littérature de l'informatique musicale. Cette section reprend les définitions du nom *geste* et précise le sens qui lui est donné dans le reste du manuscrit. Dans son usage usuel, le mot *geste* désigne un mouvement extérieur du corps, ayant une signification perceptible. Ainsi, en reprenant l'exemple du début, les *gestes* effectués par David Grimal sont notamment ceux de sa main gauche établissant la hauteur des notes ou ceux de son bras droit lui permettant de mener son archet sur la corde. L'ensemble des *gestes* en rapport avec une performance musicale sont appelés *gestes* musicaux, parmi lesquels les *gestes* de production sonore, au centre des études menées dans cette thèse.

1.1.1 Nommer les gestes musicaux : état de l'art

Plusieurs travaux ont proposé une classification des *gestes* en musique que le lecteur pourra retrouver dans les travaux (Cadoz, 1988; Wanderley et al., 2005; Dahl and Friberg, 2004; Dahl et al., 2008). Ces classifications s'appuient notamment sur la fonction du geste musical au sein de la performance. Malgré quelques différences dues à différents choix de granularité de classification, toutes s'accordent à distinguer chez un interprète les gestes *accompagnateurs* d'une part et les gestes de *production sonore* d'autre part. Contrairement aux gestes de production sonore, les gestes accompagnateurs ou *ancillaires* sont les mouvements de l'interprète faisant partie de la performance mais qui ne sont pas produits pour générer un son (Wanderley et al., 2005). Dans le cas du violon, les gestes accompagnateurs correspondent par exemple aux mouvements du violon, ou au balancement du corps du violoniste. Un raffinement des deux catégories précédentes peut être trouvé dans (Dahl and Friberg, 2004), où les auteurs scindent les gestes non reliés à la production sonore en trois :

- les gestes de communication, établissant un lien entre interprètes, chef d'orchestre et le public,

- les gestes facilitant la production du son, non reliés à la génération de son mais la facilitant,
- les gestes purement dédiés à l’accompagnement du son, mouvements que les interprètes ou l’auditoire exécutent en réponse au son produit.

Comme souligné par les auteurs, ces catégories ne sont pas disjointes : en particulier, des gestes initialement dédiés à la production sonore peuvent également contribuer à la communication entre interprètes ou entre les interprètes et le public.

1.1.2 Gestes de *production* et contrôle de l’archet

Comme leur nom l’indique, les gestes de *production sonore* ou gestes *instrumentaux* sont les mouvements de l’interprète directement reliés à la production du son. Ces gestes sont pour le musicien l’interface directe de contrôle du son de l’instrument. Dans le cas du violon, ils correspondent notamment aux mouvements du bras droit menant l’archet sur la corde. Les mouvements du bras gauche et de la main gauche modifient le son en déterminant la hauteur de la note jouée mais ne participent pas directement à la mise en vibration de la corde.

Parce qu’ils mettent en oeuvre un instrument acoustique, les gestes de production sonore sont conditionnés par les mécanismes physiques de production du son sur l’instrument. Dans le cas des cordes frottées, la production du son provient de la vibration des cordes. Cette vibration est provoquée par la friction des crins de l’archet en déplacement à une certaine vitesse (Cremer, 1984). Les mouvements de l’archet sont ainsi directement responsables de la production du son dans les cordes frottées et en conséquence font la synthèse des gestes de production sonore effectués par l’instrumentiste. Une description plus détaillée du mécanisme de production sonore chez les instruments à cordes est donnée dans le deuxième chapitre.

D’autre part, parce qu’ils sont effectués par un être humain, les gestes de production sonore sont également conditionnés par les mécanismes physiques du corps humain. Les mouvements de l’archet résultent en particulier de la coordination du bras droit : la main, l’avant-bras et le bras doivent

par exemple respecter une certaine coordination pour que l'archet ait un déplacement perpendiculaire aux cordes. Les différentes étapes d'un mouvement *tirer* suivi d'un *pousser* sont visibles sur l'illustration 1.1. Au début d'un *tirer* (mouvement de l'archet du talon à la pointe), le violoniste commence au talon de l'archet, avec son bras droit replié en forme de triangle et son poignet relevé. Ensuite, au milieu de l'archet, l'avant-bras et le bras forment un angle droit et le poignet est dans le prolongement de l'avant-bras. Finalement, à la pointe, le bras est tendu et le poignet est creusé. Pour un *pousser*, le violoniste effectue le mouvement inverse. Pour manipuler leur archet, les instrumentistes ont développé à travers un entraînement et une pratique régulière des mouvements qualifiés d'*experts*, aux caractéristiques physiques particulières (Nelson, 1983). Les mouvements possibles des instrumentistes à cordes pour produire un son sont ainsi soumis à des contraintes biomécaniques qui en conséquence conditionnent le contrôle de l'archet des musiciens.

En résumé, les gestes de production sonores, en relation directe avec le contrôle des sons émis par un instrument acoustique, se concentrent chez les cordes frottées dans les mouvements de l'archet, contrôlés par les instrumentistes en fonction de leurs capacités biomécaniques et des possibilités acoustiques de l'instrument. Dans la suite du manuscrit, le concept de gestes de production sonore est représenté par les termes plus explicites *mouvements* et *contrôle* de l'archet.

1.2 De l'intérêt de modéliser le geste

Cette section présente les problématiques de recherche les plus significatives autour du geste musical et situe le contexte général dans lequel s'inscrivent les travaux de cette thèse.

1.2.1 A la recherche de la corporalité perdue

Le foisonnement des nouvelles interfaces de contrôle gestuel et la part grandissante attribuée au domaine du contrôle et de la performance lors

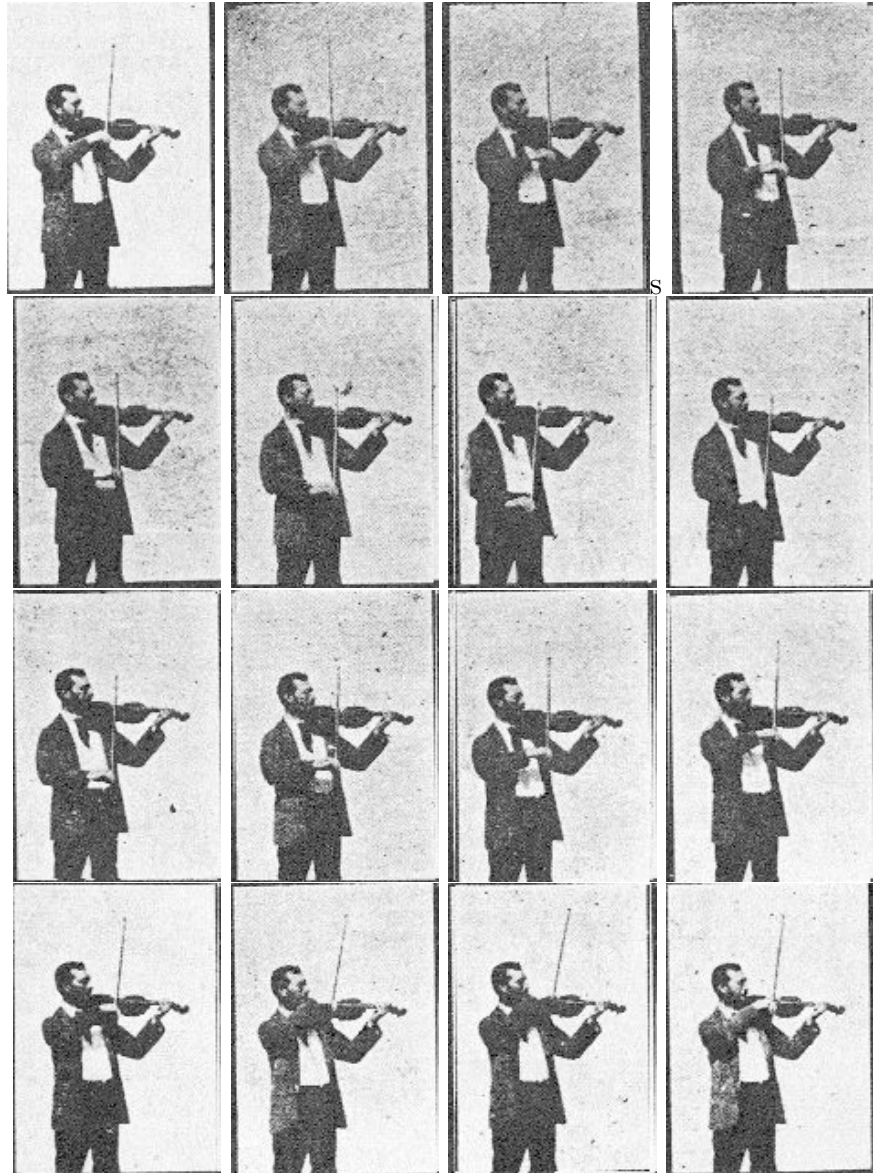


FIG. 1.1 – Animation le violoniste. G. Demeny.

des conférences en informatique et en acoustique musicales (ICMC, DAFx, NIME, ICA, ISMA, entre autres) montrent à la fois qu'il est de plus en plus aisé de réaliser de nouveaux contrôleurs originaux et qu'il y a une réelle motivation pour développer une réflexion approfondie sur les enjeux et conséquences de ces travaux en informatique musicale. Les problématiques abordées concernent notamment :

- la conception et la réalisation de nouvelles interfaces de contrôle gestuel (Waiswicz, 1985; Jordà et al., 2005; Tanaka and Knapp, 2002; Goto, 1999; Koehly et al., 2006),
- la mise en correspondance entre données gestuelles et données sonores (*mapping*) et le contrôle de la synthèse sonore (Hunt et al., 2003; Hunt and Wanderley, 2002; Arfib et al., 2002; Verfaillie et al., 2006b; Van Nort et al., 2004; Dobrian and Bevilacqua, 2003; Rank, 1999; Serafin et al., 1999),
- l'étude de critères de comparaison entre différentes interfaces de contrôle (Orio et al., 2001; Malloch et al., 2006),
- l'influence des gestes musicaux (notamment de production sonore) sur les caractéristiques timbrales du son (Schoonderwaldt et al., 2003, 2007; Guettler et al., 2003; Askenfelt, 1993; Barthet et al., 2006),
- l'influence des gestes musicaux sur la perception visuelle et auditive (Dahl and Friberg, 2007; Wanderley et al., 2005; Traube and Depalle, 2004),
- l'analyse et la synthèse de mouvements expressifs (Camurri et al., 2005; Mancini et al., 2007)
- l'étude des particularités motrices des gestes musicaux (Palmer, 2001; Baader et al., 2005; Turner-Stokes and Reid, 1998; Winold et al., 1994),
- l'étude du mouvements d'auditeurs écoutant de la musique (Styns et al., 2006),

Cette liste n'a pas vocation à être exhaustive mais cherche plutôt à montrer l'étendue du spectre d'investigations menées autour de l'étude du corps et du mouvement dans la performance musicale. Ainsi, les problématiques traitant du geste musical s'inscrivent dans différentes disciplines telles que l'acoustique, le traitement du signal et la biomécanique, domaines sur les-

quels reposent les travaux présentés dans ce manuscrit.

Conscients du bénéfice de rassembler les efforts de recherche abordant ces différentes problématiques, des chercheurs, ainsi que des musiciens, se sont rassemblés autour de différentes actions parmi lesquelles notamment l'action cost ConGAS¹ (*Gesture CONtrolled Audio Systems*) concernant l'analyse des gestes musicaux (fin en juin 2007), ou le groupe d'investigation dédié au contrôle de sons électroniques par des instruments à cordes *HotString group* (Poepel and Marx, 2007). Des efforts sont également mis sur la définition de formats pour faciliter l'échange entre différents chercheurs de données et d'analyses relatives au mouvement (Jensenius et al., 2006; Young and Deshmane, 2007).

1.2.2 La norme MIDI : séparation historique du geste et du son

La norme MIDI est reconnue dans le monde de l'informatique musicale pour sa souplesse d'utilisation et est largement utilisée (Letz, 1998). Il peut être alors intéressant dans le cadre de la recherche menée dans cette thèse de s'interroger sur les possibles raisons de ce succès. Deux aspects apparaissent notamment. La spécification de la norme MIDI permet notamment de séparer un instrument de musique en une partie contrôle et une partie sonore. Les synthétiseurs, et par la suite les contrôleurs MIDI, ont bénéficié de cette souplesse permettant ainsi de contrôler à l'aide d'une même interface, en l'occurrence un clavier, plusieurs sons différents et de natures différentes : sons synthétiques modulables, sons échantillonnés d'autres pianos (Steinway, Bösendorfer) ou d'autres instruments (guitare, glockenspiel, trompette, violon). Cette dissociation a fait ainsi de MIDI l'une des premières solutions permettant la création d'instruments de musique numériques (*Digital Music Instruments* ou *DMI*) (Wanderley and Depalle, 2004).

Le deuxième aspect concerne la manière dont peuvent être effectuées les différentes associations entre la partie contrôle et la partie sonore. La norme

¹<http://www.cost287.org/>

MIDI permet la description des actions d'un instrumentiste jouant sur le clavier : les événements NOTE-ON, NOTE-OFF modélisent les instants de pression et de relâchement de la touche, le paramètre VELOCITY reflète l'impulsion donnée par le pianiste à l'enfoncement de la touche. Cette description élémentaire fait ainsi de la norme MIDI l'une des premières *modélisations* des gestes d'un instrumentiste, en l'occurrence ceux du pianiste (Letz, 1998). Par cette modélisation, la norme MIDI impose des conditions sur les liens entre les parties contrôle et génération de sons, permettant d'opérer les associations et réassociations multiples décrites auparavant tout en conservant la pertinence du jeu du clavier. Cet aspect, bien que moins fréquemment souligné que le premier, a sans doute tout aussi substantiellement contribué au succès de la norme MIDI.

1.2.3 Le cas des instruments à son entretenu

Cependant, cette norme, héritée des années 80 et de l'incroyable développement des synthétiseurs à ce moment, subit de nombreuses critiques (Moore, 1988; McMillen, 1994; Wessel and Wright, 2002). Notamment, si le contrôle est convaincant lorsqu'il s'agit de sons frappés comme le piano ou le glockenspiel, il l'est beaucoup moins pour des sons entretenus comme dans les cordes frottées. Bien qu'elle soit très utilisée dans les synthétiseurs actuels, l'association entre un contrôle au clavier et un son entretenu, c'est-à-dire entre les événements discrets autorisés par la norme MIDI et le contrôle continu inhérent aux instruments à son entretenu, est particulièrement inadéquate (McMillen, 1994; Wessel and Wright, 2002; Rank, 1999).

Les solutions offertes dans les articles (McMillen, 1994) et (Wessel and Wright, 2002), résolvent ce problème en visant les mêmes objectifs : augmenter la bande passante et améliorer la résolution temporelle, jugées trop limitées. Bien qu'apportant de réelles améliorations, ces travaux n'abordent pas un troisième aspect présent dans la norme MIDI : la modélisation des gestes instrumentaux. Les travaux de Chafe (1989) et Rank (1999) présentent les premiers essais de modélisation des paramètres de contrôle pour l'excitation d'une corde frottée, modélisant ainsi les actions de l'instrumentiste,

mais tous deux relèvent la difficulté de définir des évolutions temporelles pertinentes pour ces paramètres.

Les instruments à sons entretenus offrent effectivement de grandes possibilités en terme de *modes de jeu*. En faisant varier les paramètres de fonctionnement de leurs instruments, les musiciens arrivent ainsi à produire des différences de timbre et des dynamiques temporelles particulières. Dans le cas des instruments à cordes frottées, les interprètes peuvent par exemple jouer sur la touche et produire un son flûté (*flautendo*) ou encore jouer avec le bois de la baguette pour obtenir un son bruité (*con legno tratto*). En contrôlant la dynamique de l'archet, ils peuvent produire différentes techniques d'archet telles que *Détaché*, *Martelé* et *Spiccato* qu'ils utilisent pour construire leurs interprétations. Ces modes de jeu, développés par plusieurs générations d'instrumentistes, reflètent les capacités expressives de l'instrument et présentent ainsi un intérêt non négligeable pour l'étude des gestes instrumentaux.

1.2.4 Musique mixte et interaction musicale à l'IRCAM

L'une des missions de l'IRCAM est d'accorder une place importante à la création et à la promotion de musiques nouvelles, aux frontières entre art et sciences. Cette ambition en a fait l'un des acteurs importants de l'histoire de la musique contemporaine, comme en atteste notamment ses contributions à la musique mixte². Dans les années 50, époque des premières pièces mixtes³, la partie électronique était fixée sur une bande magnétique : la performance de la pièce demandait alors aux musiciens l'effort de suivre le déroulement de la bande, ce qui ne leur laissait que peu ou pas de possibilités pour développer une interprétation de la pièce. Bien que limitatives sur ces aspects, ces pièces

²dans cette thèse, nous adoptons une définition générale désignant comme musique mixte toute oeuvre de musique impliquant un interprète humain jouant un instrument de musique acoustique et des sons électroniques, échantillonnés, ou traités, "joués" par une machine

³*Orphée* (Pierre Schaeffer, 1951), *Musica su due dimensioni* (Bruno Maderna, 1957-8), *Déserts* (Edgard Varèse, 1954).

sont les premiers témoins de la volonté des compositeurs à intégrer une machine comme actrice de l'oeuvre musicale au même titre qu'un interprète vivant.

Par la suite, chercheurs et compositeurs se sont interrogés sur les possibilités de créer des systèmes propices à une véritable interactivité musicale entre un humain et un ordinateur, idéalement du même ordre que celle reliant différents interprètes jouant ensemble, où le musicien et la machine peuvent s'écouter et se répondre. Le suivi de partition (Vercoe, 1984; Dannenberg, 1984; Raphael, 1999; Orio and Dechelle, 2001; Schwarz et al., 2004; Cont, 2008) permet ainsi de définir une partition propre à un ordinateur, rejouée en suivant le déroulement de la partition de l'interprète. L'instrument acoustique suivi peut être un dispositif de contrôle MIDI (clavier muet), un instrument "midifié" par un système de capteurs mécaniques (piano Steinway⁴, flûte MIDI⁵) ou acoustiques (micro-contacts⁶), ou également un instrument traditionnel sans aucun capteur auquel cas le signal acoustique capté par un micro est suivi⁷. En rendant la machine à l'écoute du musicien, ce système permet à l'instrumentiste de retrouver une grande liberté d'interprétation.

Une approche complémentaire consiste à s'appuyer sur les formes et structures musicales d'une performance réelle pour proposer un système d'interaction agissant au niveau du style musical de l'interprète (Assayag et al., 1999, 2006; Rowe, 1992; Pachet, 2002). Cette approche offre notamment la possibilité à un interprète d'improviser avec son propre matériau, déstructuré et restructuré par la machine. Ce type d'interactions met généralement en oeuvre deux phases, une première où la machine écoute et une deuxième où elle répond, et établit ainsi un véritable dialogue entre le musicien et une machine.

Les travaux de Risset and Van Duyne (1996) présentent encore un angle

⁴*Pluton* (Philippe Manoury, 1988), piano MIDI.

⁵...*Explosante fixe*... (Pierre Boulez, 1993), initialement conçue pour flûte MIDI elle a été reprise en 2006 en version suivi audio.

⁶*Neptune* (Philippe Manoury, 1991), vibraphones MIDI, tam-tam et marimba.

⁷*En écho* (Philippe Manoury, 1993) pour soprano, *Jupiter* (Philippe Manoury, 1987) pour flûte, *Anthème II* (Pierre Boulez, 1997) pour violon et plus récemment ...*of Silence* (Marco Stroppa, 2007) pour saxophone.

différent. L'interaction musicale se place au niveau des gestes instrumentaux d'un pianiste, transformant ainsi la machine en un miroir du musicien : la machine écoute les gestes de l'instrumentiste, exprimés sous la forme de messages MIDI (NOTE-ON, NOTE-OFF, VELOCITY) et répond sous la forme de nouveaux messages MIDI. Ce système donne alors place à un véritable duo pour un seul pianiste⁸.

Dans ce contexte général d'interactions musicales, la question des gestes musicaux, au sens précédemment défini, a des résonances particulières. La considération des mouvements des musiciens permet d'ajouter une dimension supplémentaire à la réflexion des chercheurs et des compositeurs qui se trouvent dès lors en mesure de définir des systèmes d'interaction établissant un lien physique, "incarné", entre musiciens et machines. Un tel lien apporterait des éléments de réponse substantiels à la question de la dissociation entre expression artistique et technologies posée en introduction. Il est par ailleurs à noter que cette dimension se présente déjà de manière plus ou moins latente dans les travaux précédents, notamment au travers des instruments midifiés mais également avec les analyses audio (cf chapitre 2, mesure des gestes). L'étude du mouvement des musiciens contribue néanmoins explicitement dans ce sens. Les travaux menés dans cette thèse s'attachent en particulier aux gestes instrumentaux expressifs dans le cas d'instruments à sons entretenus. Ce domaine relativement peu étudié nécessite d'être d'avantage exploré afin de permettre la définition telles interactions.

1.3 Problématique de la thèse

Un instrument acoustique émet un son sous certaines conditions précises d'excitation. Pour jouer d'un instrument à cordes, le musicien doit en premier lieu effectuer les mouvements nécessaires pour déclencher et entretenir la vibration des cordes de son instrument. Ces mouvements sont de plus conditionnés par les possibilités biomécaniques du corps du musicien. L'ensemble des contraintes acoustiques et biomécaniques définissent alors un champ des

⁸*Duo pour un pianiste*, (Risset, 1989).

possibles dans lequel différentes stratégies d'action peuvent exister et que les instrumentistes utilisent pour réaliser différents modes de jeu expressifs (Bril, 2002; Goasdoué, 2002).

En se concentrant sur des mesures d'instrumentistes en situation de jeu, cette thèse propose d'explorer et tente de caractériser ces espaces de possibles. Dans ce contexte, l'étude de différentes techniques d'archet, exploitant ces espaces, peut permettre de déterminer de telles caractéristiques. L'idée n'est alors pas de trouver une modélisation pour chaque mode de jeu étudié mais plutôt d'en dégager des principes sous-jacents plus généraux. Les travaux présentés dans ce manuscrit se concentrent ainsi sur un ensemble de trois modes de jeu (*Détaché*, *Martelé* et *Spiccato*) choisis au préalable pour leur différences et complémentarités. De même, l'idée n'est pas de mener des statistiques sur un grand nombre d'instrumentistes, dont la variabilité risque d'être élevée (Dahl, 2000; Winold et al., 1994), mais plutôt d'étudier en profondeur les stratégies possibles utilisées par un nombre plus restreint d'instrumentistes experts. L'étude des données du mouvement d'instrumentistes en jeu réel implique également la donnée d'une segmentation temporelle définissant des unités d'analyses à comparer. Différentes approches sont proposées dans cette thèse.

Cette thèse présente quatre articles abordant cette problématique. Un résumé des contributions est proposé dans le chapitre II et une discussion générale mettant les résultats en perspective est menée dans le chapitre III.

Chapitre 2

Contribution

Les articles constituant ce manuscrit ont pour objet l'étude du contrôle de l'archet des instrumentistes à cordes en situation de jeu. La compréhension de différentes stratégies utilisées dans la maîtrise de l'archet, notamment celles ayant trait à des fins expressives, peut effectivement apporter des éléments fondamentaux pour le contrôle de synthèse sonore ainsi que pour l'éducation musicale. Pendant leur apprentissage, les instrumentistes à cordes acquièrent plusieurs techniques d'archet élaborant ainsi leur maîtrise instrumentale à partir de laquelle ils seront capables de formuler et d'articuler un discours musical. L'analyse de techniques d'archet standards, comme le *Détaché* et le *Martelé*, permet alors d'étudier cette maîtrise en la décomposant sur ces formes prototypiques. En faisant varier le contexte musical (nuances, tempo, accentuations, modes de jeu) il est possible de créer diverses situations d'analyse des stratégies de contrôle des instrumentistes (Winold et al., 1994; Wanderley, 2001; Palmer, 2001; Dahl, 2005; Goebel et al., 2005).

Les travaux menés dans cette thèse s'intéressent à l'étude et à la modélisation des mouvements de l'archet et du rapport entre ces mouvements et le son dans l'exécution de différents modes de jeu standards des cordes frottées. Ils s'appuient notamment sur l'hypothèse que dans le cadre des cordes frottées, les mouvements des musiciens sont conditionnés par deux contraintes en particulier : d'une part une contrainte acoustique consistant à la combinaison des paramètres de jeu permettant la mise en vibration de

la corde (Askenfelt, 1989; Woodhouse and Galluzzo, 2004; Guettler, 2002b; Guettler and Askenfelt, 1997) et d'autre part, une contrainte biomécanique relative aux capacités physiques de l'instrumentiste (Nelson, 1983; Hogan, 1984; Hogan and Sternad, 2007). Ces deux contraintes définissent un champ des possibles dans lequel différentes stratégies d'action peuvent exister et que les instrumentistes utilisent pour réaliser différents modes de jeu expressifs (Goasdoué, 2002). Ainsi, les analyses présentées ici traitent la problématique du contrôle de l'archet des instrumentistes à cordes en tirant parti des points de vue complémentaires de ces disciplines.

- **L'article I** tente d'explorer les domaines de possibles dans l'espace des paramètres de jeu en s'intéressant notamment à l'invariabilité et la variance de l'accélération de l'archet sur chaque coup en fonction des modes de jeu *Détaché*, *Martelé* et *Spiccato*, de la nuance et du tempo, joués par plusieurs instrumentistes.
- **L'article II** définit des unités d'analyse centrées sur les transitions entre coups d'archet. Il propose une vue plus détaillée des modes de jeu *Détaché* et *Martelé* en analysant les rapports entre geste et son, i.e. entre accélération de l'archet et vibration de la corde, pendant la production de différentes qualités d'articulations entre notes. Il souligne ainsi le fait que la maîtrise des transitions forme une partie constitutive du contrôle de l'archet.
- **L'article III** approfondit la notion de domaines de possibles en étudiant plus particulièrement l'influence du tempo sur les mouvements d'archet d'instrumentistes effectuant un *accelerando* / *decelerando*.
- **L'article IV** met en avant l'influence de comportements d'anticipation et de "coarticulation gestuelle" dans l'exécution de modes de jeu en proposant des modèles du mouvement inspirés de travaux en biomécanique.

2.1 Mesures des gestes du violoniste en situation de jeu : méthodologie

Les articles présentés dans ce manuscrit s'appuient tous sur la mesure du mouvement d'instrumentistes en situation de jeu. La méthodologie suivie est décrite dans cette section. Les idées et hypothèses générales soutenant les protocoles de mesure sont présentées ainsi que les divers systèmes de captations du mouvement utilisés.

2.1.1 Principaux paramètres physiques des cordes frottées

Les travaux de cette thèse s'intéressent aux gestes de production sonore, c'est-à-dire aux variables physiques responsables de la mise en vibration de la corde. Les principaux paramètres gouvernant le déplacement de la corde dans un instrument à corde frottée sont notamment décrits dans (Askenfelt, 1986, 1989) et dans (Guettler, 2002a). Ceux-ci comprennent la position transversale de l'archet (perpendiculairement aux cordes), ses dérivées la vitesse et l'accélération transverses, la force verticale (ou pression) de l'archet sur les cordes et la position de l'archet sur la corde. Les influences respectives de chacun de ses paramètres sur le mouvement de la corde ne sont pas indépendantes. Seulement certaines combinaisons de ses paramètres permettent de placer la corde dans un mouvement périodique, donnant ainsi la possibilité d'une hauteur perçue.

Dans sa description la plus acceptée, ce régime périodique se caractérise par la vibration de la corde en forme de V, appelé coin de Helmholtz. Celui-ci se déplace le long de la corde lorsque l'archet la frotte dans des conditions optimales créant ainsi un régime de Helmholtz (Woodhouse and Galluzzo, 2004). A noter qu'il existe d'autres descriptions, moins répandues, que le lecteur peut retrouver dans l'article de McIntyre et al. (1984). S'appuyant sur la description de Helmholtz, le diagramme de Schelleng (Schelleng, 1973) établit un triangle de fonctionnement permettant l'entretien d'un mouvement (ou *régime*) de Helmholtz en fonction de combinaisons entre la pression d'archet

et la position sur la corde. De nombreux travaux en acoustique s'intéressent à déterminer l'influence de chacun des principaux paramètres de l'archet sur le déplacement de la corde et sur le spectre du son obtenu, notamment (Guettler, 2004b; Guettler et al., 2003; Schoonderwaldt et al., 2003, 2007; Woodhouse and Galluzzo, 2004; Askenfelt, 1989).

L'établissement d'un régime de Helmholtz est également soumis à des contraintes précises de manière à être "musicalement acceptable"¹. Il existe différents types de régimes transitoires pendant lesquels se crée un régime périodique (Guettler and Askenfelt, 1997). A l'aide d'une machine contrôlant l'archet sur une corde (*bowing machine*), trois régimes différents ont été répertoriés par Guettler et Askenfelt en fonction de la pression d'archet et de l'accélération initiale pour des mises en vibration initiale de la corde, régimes dits *pre-Helmholtz*. Ils se traduisent par une attaque "étouffée - écrasée" (choked - creaky) "détendue - granuleuse" (loose - scratchy) ou "parfaite" (perfect) selon les termes utilisés par les auteurs. Des études psychoacoustiques menées par les mêmes auteurs révèlent également les limites acceptables de durée de ces régimes *pre-Helmholtz* (Guettler and Askenfelt, 1997) : la durée doit être inférieure à $50ms - 90ms$ selon le type de vibrations *pre-Helmholtz*.

Les études acoustiques présentées permettent d'identifier les paramètres importants et de quantifier d'un point de vue théorique leur rôle dans le contrôle de la vibration de la corde. De manière à étudier les différentes stratégies de contrôle d'instrumentistes réels, il est donc nécessaire de pouvoir les enregistrer en situation de jeu. Divers systèmes de mesure sont aujourd'hui à disposition pour y arriver. L'une des principales conditions pour réaliser des études en situations réelles est d'avoir un système permettant la mesure précise des variations des paramètres présentés, tout en minimisant l'intrusion du système dans les mouvements des instrumentistes.

¹d'après le jugement émis par 20 instrumentistes à cordes professionnels et étudiants.

2.1.2 Capture des paramètres de l'archet

Plusieurs systèmes permettant l'acquisition des mouvements de l'archet des instrumentistes existent. Sont décrits ici les plus significatifs ainsi que ceux utilisés dans les travaux de cette thèse.

Systèmes de capture du mouvement

L'exemple le plus représentatif de ce type de captation est le système *VI-CON*. Il repose sur le suivi (*tracking*) de marqueurs par un réseau de caméras rapides à haute résolution. Les marqueurs sont des réflecteurs passifs, non-filaires, et sont en conséquence peu invasifs. En plaçant des réflecteurs sur un sujet, dans le cas présent sur l'archet et sur l'instrument, il est possible de mesurer des déplacements avec une précision spatiale en dessous de $1mm$, à une fréquence pouvant atteindre $500Hz$ dans un volume de l'ordre de $1m^3$. Cette configuration a par exemple été utilisée pour récolter les données utilisées dans les articles III et IV. Les marqueurs utilisés sont généralement sphériques ou demi-sphériques (il existe également des gommettes plan) et leur taille est variable selon l'utilisation envisagée : de $5mm$ à plusieurs centimètres de diamètre. Le placement des marqueurs utilisé pour les articles III et IV sont visibles sur la figure 2.1.

Toutefois, ce système présente quelques inconvénients. Malgré sa souplesse d'utilisation, le réseau de caméra est relativement encombrant : il est contraignant à déplacer et le système reste généralement attaché à une salle dédiée. De plus, de part la technologie employée, ce type de système de captation est relativement onéreux. Cependant, malgré ces inconvénients, ce système est l'outil idéal pour effectuer des mesures précises tout en minimisant la gêne pour l'instrumentiste.

Ce système a précédemment été employé sur des violoncellistes par (Turner-Stokes and Reid, 1998) pour explorer le potentiel des analyses 3-D afin de détecter d'éventuels problèmes musculaires chez les musiciens. A l'aide d'un système similaire (*ELITE*), les auteurs de (Baader et al., 2005) ont étudié la coordination bimanuelle chez six violonistes en mesurant simultanément les déplacements des doigts de la main gauche (hauteur des notes) et du bras



FIG. 2.1 – (gauche) : réseau de caméras rapides enregistrant un violoniste. (droite) : Placement des réflecteurs utilisé pour mesurer les mouvement du violon et de l'archet.

droit (contrôle de l'archet).

Il existe d'autres systèmes de capture de mouvement, notamment s'appuyant sur des champs électromagnétiques pour mesurer les déplacements. Les systèmes *Polhemus* reposent sur des marqueurs actifs placés dans un champ électro-magnétique. Chaque marqueur est identifié et suivi en s'appuyant sur les perturbations qu'il crée dans le champ électro-magnétique. Cette technologie évite ainsi les éventuelles faiblesses d'un suivi vidéo pouvant perdre ou intervertir certains marqueurs lors de l'enregistrement. Toutefois, l'alimentation des marqueurs actifs conditionne leur encombrement. Les marqueurs sans fils fonctionnent sur piles et sont en conséquence relativement volumineux et lourds. Les marqueurs de taille et poids plus faibles sont en contre-partie filaires. Le système *Polhemus* a précédemment été utilisé pour mesurer les gestes de violonistes dans (Peiper et al., 2003; Maestre et al., 2007) et un système similaire (*Ascension SpacePad*) a été utilisé dans (Goudeseune et al., 2001).

Capteurs embarqués sur l'archet

L'utilisation de capteurs sur un archet permet la mesure des paramètres de jeu du violon de manière fine et généralement à moindre coût. L'archet ainsi modifié autorise notamment une certaine souplesse d'utilisation : le même système peut être utilisé pour effectuer des mesures en différents endroits et/ou pour une performance musicale. Les premières mesures détaillées des paramètres de jeu des cordes frottées ont été menées par Askenfelt (Askenfelt, 1986, 1989). Un archet de violon a été équipé de jauges de contraintes pour mesurer la force de l'archet sur les cordes. La position transversale et la position de l'archet au chevalet ont été déduites par un système de courant dissipé par une bande résistive. Pour la position transversale, une bande résistive est insérée le long de la mèche de l'archet, dans le cas de la position de l'archet au chevalet, la corde est utilisée comme résistance. Les positions sont alors directement corrélées à l'intensité du courant en fonction des propriétés résistives des matériaux : si les résistances ont un comportement linéaire alors les positions sont proportionnelles à l'intensité du courant. Ce

dispositif est l'un des premiers à avoir permis la mesure de paramètres de l'archet en situation de jeu.

Le groupe de recherche *HyperInstrument* (Machover and Chung, 1989) du M.I.T., fondé en 1989 par Tod Machover, a également contribué de manière substantielle au développement d'instruments à cordes modifiés, par exemple avec l'*HyperCello*². Young a notamment oeuvré à l'implémentation d'un archet sans fil (Young, 2003), le *HyperBow*, dédié à la performance artistique et utilisé dans *ToySymphony*, opéra de Tod Machover, et dans (Young et al., 2006). Ce système repose sur l'utilisation d'accéléromètres et de gyroscopes pour la mesure de l'accélération de l'archet et de son angle avec les cordes, ainsi que d'un système de couplage capacitif entre une bande résistive collée à l'archet et une antenne fixée derrière le chevalet pour la mesure des positions transverse et entre archet et chevalet. De plus, Young incruste des jauges de contraintes dans la baguette de l'archet de manière à mesurer sa déformation. Dans sa thèse (Young, 2007), Young a réalisé un travail de calibration de son système permettant d'utiliser l'*HyperBow* pour des mesures quantitatives : l'exactitude des mesures données par son système est évaluée et les signaux sont convertis en unité S.I. Toutefois à ce jour, son manuscrit n'est pas disponible et il n'est malheureusement pas possible de connaître les résultats détaillés de son évaluation.

Plusieurs versions de modules de captation sans fil ont été réalisés à l'IR-CAM dans le contexte du projet du *violon augmenté*. Comme le *HyperBow*, ces modules embarquent des accéléromètres. En revanche, contrairement à ce dernier, les modules ont été conçus pour pouvoir être placés sur tous les archets sans nécessiter d'opérations lourdes ou irréversibles (collages entre autres). Ainsi les mesures possibles avec ces modules concernent essentiellement les dynamiques de l'archet. La captation de la pression de l'archet sur les cordes a été développée séparément et est traitée dans la sous-section dédiée. Les modules de la figure 2.2 représentent deux versions successives des systèmes. Ils se glissent sur la hausse de l'archet à l'aide d'une bague en carbone et mesure les accélérations de l'archet selon 3 axes : dans le sens de

²<http://brainop.media.mit.edu/Archive/Hyperinstruments/hypercello.html>

la baguette, parallèlement aux cordes et orthogonalement aux deux autres directions. Ils représentent un poids total (avec batteries) de 20g et 17g respectivement. Le module représenté sur la figure 2.3 est la dernière version en date. Il comprend un capteur mesurant l'accélération dans 3 directions et 2 gyroscopes donnant accès à 5 degrés de liberté (DoF). Le capteur est glissé dans une "chaussette" spécialement conçue pour la hausse de l'archet, ne lui rajoutant ainsi que quelques grammes. Le capteur est relié à un émetteur sans fil fixé sur un bracelet que doit porter l'instrumentiste. Ce système est notamment utilisé dans le cadre d'une création musicale pour quatuor à cordes *augmenté* de la compositrice Florence Baschet et du projet européen I-MAESTRO étudiant l'utilisation de nouvelles technologies dans la pédagogie musicale.

Il existe de nombreux autres systèmes utilisant des instruments à cordes modifiés parmi lesquels notamment les créations de Trueman and Cook (1999), de Nichols (2002), de Goto (1999), de Overholt (2005a), de Freed et al. (2006), de Young and Serafin (2007), et de Chafe³ dont la plupart a été conçue à dessein artistique. Il convient aussi de mentionner les travaux originaux de Jon Rose⁴.

Combinaison capteurs-vidéo

De manière à mesurer les paramètres physiques du jeu du violon, les capteurs doivent subir une phase de calibration, permettant de corriger les éventuelles non-linéarités et d'obtenir des valeurs en unités S.I. Toutefois, certains capteurs utilisés seuls ne permettent pas d'obtenir des mesures justes : par exemple la valeur donnée par un accéléromètre dépend à la fois de son angle d'inclinaison et de l'accélération subie. Il convient parfois de combiner plusieurs capteurs entre eux, comme l'association fréquente accéléromètre - gyroscope. Cependant, sur un archet où l'encombrement est critique, une telle solution est discutable.

Une solution implémentée pendant la thèse réalise un compromis entre

³<http://news-service.stanford.edu/pr/97/970115chafeprof.html>, testé le 10/12/07

⁴<http://www.jonroseweb.com/>, testé le 10/12/07

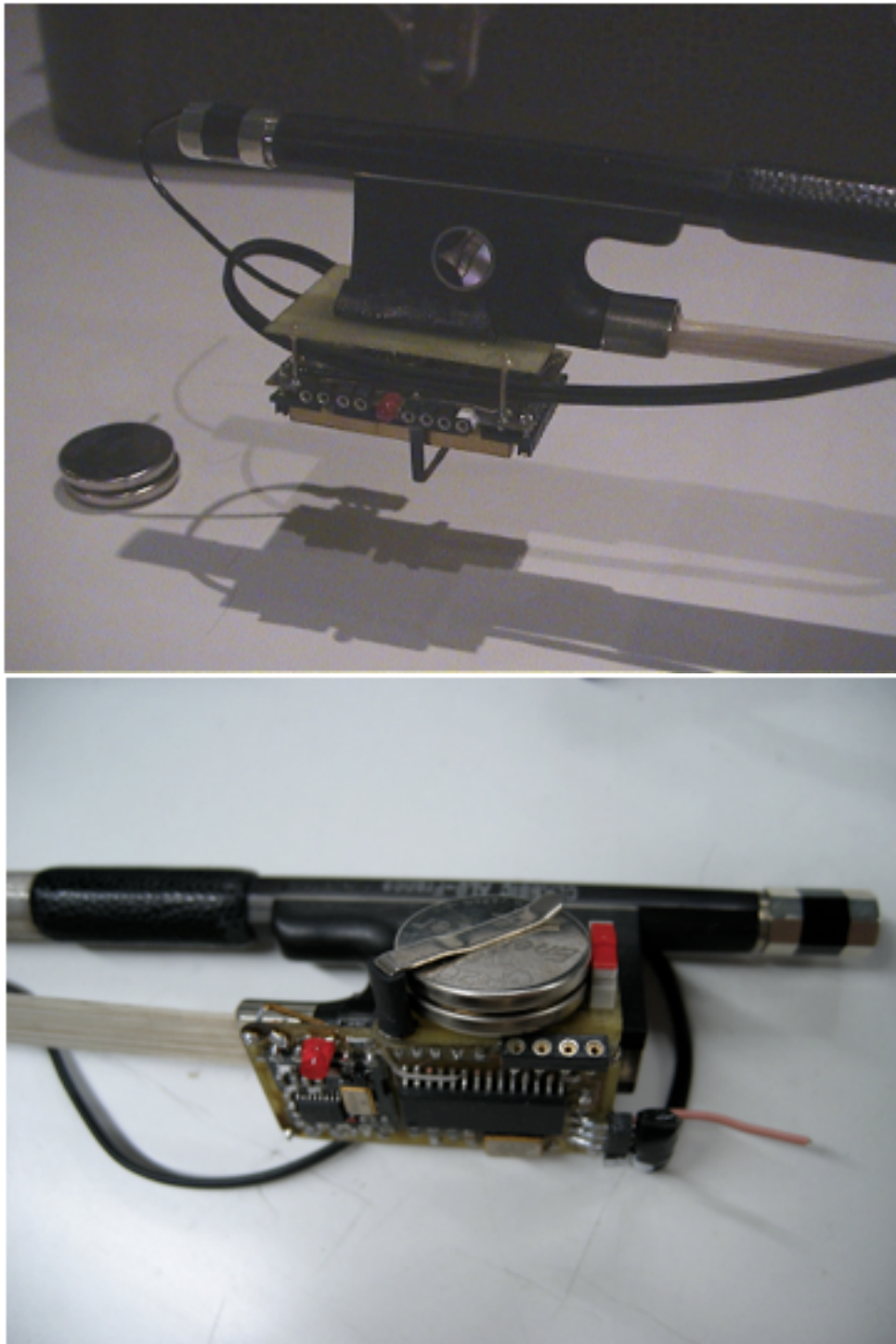


FIG. 2.2 – Première et deuxième versions des modules de captations sans fil. Dans ces versions, les accéléromètres, le microcontrôleur, l'émetteur sans fil et les batteries sont placés sur une pince en carbone glissée sur la hausse de l'archet.

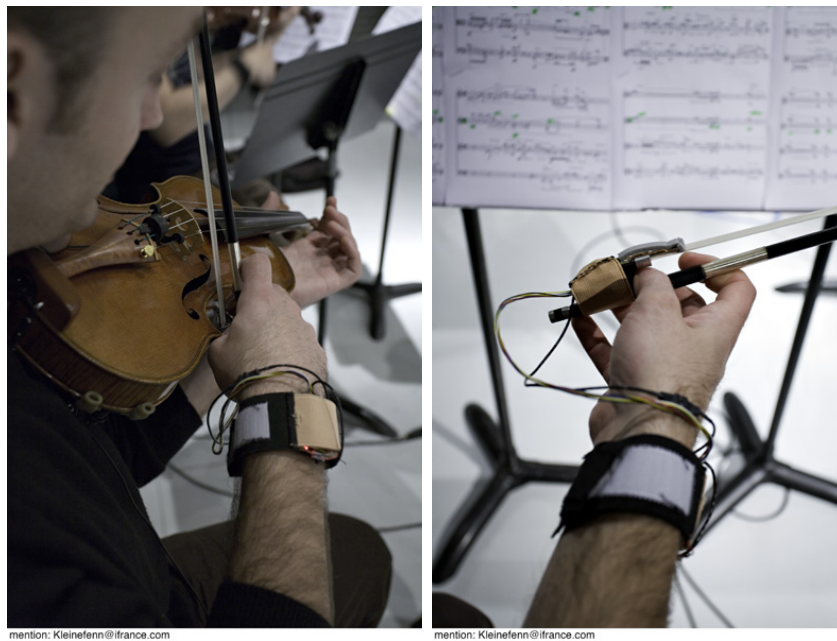


FIG. 2.3 – Dernière génération de module de captation de l'IRCAM. Seuls les accéléromètres sont placés sur l'archet et le reste de l'électronique, notamment l'émetteur sans fil et les batteries, est déporté sur le poignet. Egalement visible sur la photo de gauche placé sur la mèche de l'archet, un prototype de capteur de pression, recouvert de mousse protectrice.

encombrement sur l'archet et justesse des mesures en combinant des capteurs embarqués avec un autre type de captation, i.e. la vidéo, pour corriger les défauts des capteurs (Schoonderwaldt, Rasamimanana and Bevilacqua, 2006). Dans cet article, la vitesse de l'archet est calculée grâce à une intégration par morceaux des signaux d'accéléromètres. La sommation est réinitialisée à chaque changement de sens de l'archet, détecté à partir du suivi vidéo de marqueurs sur l'archet, permettant ainsi de minimiser l'accumulation d'erreurs. Dans les cas simples où l'on peut supposer que l'angle de l'archet reste constant sur chaque coup, les signaux d'accéléromètres ont une valeur de base constante que l'on peut soustraire à chaque réinitialisation. La vitesse de l'archet peut ainsi être reconstruite à partir des signaux des accéléromètres corrigés et bénéficier de leur résolution temporelle généralement choisie plus grande. Lorsque la désynchronisation et la gigue entre le flux d'images vidéo et les signaux d'accélérations restent faibles (respectivement $25ms$ et $40ms$), cette méthode offre une alternative intéressante et bon marché aux systèmes type *VICON* pour obtenir des profils de vitesses acceptables (coefficient de corrélation entre 0.9 et 1.0) (Schoonderwaldt, Rasamimanana and Bevilacqua, 2006). Les détails de l'implémentation et de l'évaluation sont disponibles dans l'article, disponible en annexe.

Capteur de pression de l'archet

S'il existe plusieurs manières efficaces de mesurer la dynamique de l'archet en situation de jeu, la mesure de pression de l'archet sur les cordes est un problème difficile auquel peu de solutions robustes ont été proposées. Les systèmes existants reposent sur des modifications irréversibles d'un archet comme dans (Young, 2003) où des jauges de contraintes sont collées sur la baguette, ou des opérations relativement lourdes comme dans (Askenfelt, 1986, 1989) où la mise en place des capteurs implique le démêchage et remêchage de l'archet. Pendant le temps de cette thèse et dans le cadre de la sienne, Matthias Demoucron a notamment conçu et développé un capteur de pression d'archet précis, non destructif pour l'archet et avec un système de fixation ergonomique (Demoucron et al., 2006; Demoucron, 2007). Le prin-

cipe repose sur une jauge de contrainte fixée à une plaque de métal en appui sur la mèche de l'archet et permettant d'en mesurer les déflexions : il est ainsi possible de déduire la pression de l'archet sur la corde à partir de ses informations. De premiers tests montrent que les mesures fournies par ce capteur sont particulièrement précises et robustes. Ce capteur prometteur est visible sur la figure 2.4 et le lecteur est invité à se reporter à la thèse de Matthias Demoucron pour de plus amples informations.



FIG. 2.4 – Capteur de pression : un jauge de contrainte est placée sur une plaque en appui sur la mèche de l'archet. Elle mesure les déflexions de la mèche à partir de laquelle il est possible de déduire la pression sur la corde. L'ensemble est fixé sur la bague de la hausse de l'archet.

Capture des signaux acoustiques

De manière à étudier l'influence des mouvements de l'archet sur le son, une capture des signaux acoustiques est nécessaire. Cette capture peut s'effectuer

en différents endroits, en fonction des objectifs. Dans les travaux (Askenfelt, 1986, 1989; Guettler and Askenfelt, 1997; Guettler et al., 2003; Young, 2003), les mesures concernent le déplacement de la corde. Les systèmes de mesures sont alors placés en dessous de la corde et enregistrent la force exercée par la corde sur le chevalet ou la vitesse de déplacement de la corde, par exemple avec des capteurs piezo-électriques. Les propriétés de vibration de la corde peuvent alors être étudiées. Dans l'article (Leroy et al., 2006), les auteurs ont conçu et réalisé des capteurs s'appuyant sur la réflexion de diodes lasers placées sous les cordes de manière à pouvoir mesurer la hauteur des notes (*pitch*) individuellement sur chaque corde. Ce système élaboré dans le contexte du projet du *violon augmenté* permet la captation du mouvement des cordes de manière moins intrusive que les autres systèmes optiques, fonctionnant en général en transmission (Overholt, 2005a), et sur des cordes non nécessairement ferromagnétiques (e.g. cordes en boyau). Ces deux propriétés permettent ainsi l'utilisation de ces capteurs sur tous les violons.

Une autre approche consiste à utiliser des capteurs piezo-électriques enregistrant directement la vibration de la corde. Toutefois, sur un instrument acoustique, le placement de capteurs piezo-électriques implique la modification de tout ou d'une partie du chevalet : outre la réserve des musiciens quant à l'emploi de ce genre de capteurs sur leur instrument, cette approche implique en plus une modification du son, de type sourdine, pouvant être gênante pour des instrumentistes en situation de jeu. Une autre solution consiste également à utiliser un violon électrique à la place d'un instrument acoustique traditionnel en tirant partie des microphones déjà intégrés sous le chevalet. Toutefois, les sensations sur ce genre d'instrument sont significativement différentes de celles sur un instrument acoustique : notamment, sans amplification, la puissance rayonnée par un violon électrique est presque nulle car il est dépourvu d'une caisse de résonance, tandis que l'amplification restitue généralement un son sensiblement moins "mordant", ce qui dans les deux cas peut être perturbant pour un instrumentiste traditionnel.

Une dernière approche consiste à enregistrer le son émis par l'instrument à l'aide de microphones aériens. Les signaux acoustiques enregistrés tiennent alors compte des effets de filtrage de la caisse de résonance. Cette méthode

est la moins intrusive pour l'instrument et est celle qui a été le plus souvent utilisée dans la thèse. Les signaux acoustiques utilisés dans les travaux de cette thèse ont été mesurés à l'aide d'un microphone aérien DPA 4021, fixé derrière le chevalet.

Capture indirecte des gestes instrumentaux

Une approche parallèle à l'utilisation de systèmes de captation consiste à étudier la trace laissée par les gestes instrumentaux dans le signal acoustique. Les travaux suivant cette approche cherchent ainsi à ne pas avoir recours à des systèmes de captation, évitant ainsi leur inévitable encombrement, et contournant également les éventuelles difficultés de se procurer l'unique exemplaire d'un système de captation (Kapur et al., 2004b). Il est possible par exemple de retrouver à partir de caractéristiques du spectre des paramètres de jeu d'instrument tel que le point de pincage d'une corde de guitare (Orio, 1999; Traube et al., 2003) ou le doigté des notes harmoniques à la flûte (Kereliuk et al., 2007) ou encore la direction du pincement sur un sitar (Kapur et al., 2004b). Ces paramètres extraits du signal acoustique peuvent être utilisés pour contrôler des algorithmes de synthèse sonore (Hikichi et al., 2004).

En ce qui concerne les mouvements relatifs à la production sonore chez les cordes frottées, en se plaçant dans l'hypothèse d'un mouvement de Helmholtz, la vitesse de l'archet détermine l'amplitude de déplacement de la corde frottée : pendant la phase d'adhérence du cycle adhérence-glissement, la vitesse de l'archet et celle de la corde sont identiques. Comme le montre la figure 2.5, l'amplitude du signal acoustique émis est ainsi directement corrélé à la vitesse de l'archet. Toutefois, des "perturbations" existent dans cette correspondance, notamment des effets de résonance de la caisse de l'instrument et les effets du vibrato de la main gauche, ayant tous deux des influences significatives sur l'enveloppe sonore. L'article (Poepel, 2004) exploite cette capture indirecte de la vitesse de l'archet en contrôlant différents algorithmes de synthèse sonore à l'aide de l'amplitude du signal audio d'un alto et son *pitch*, faisant ainsi l'économie de dispositifs de captations supplémentaires. Cette

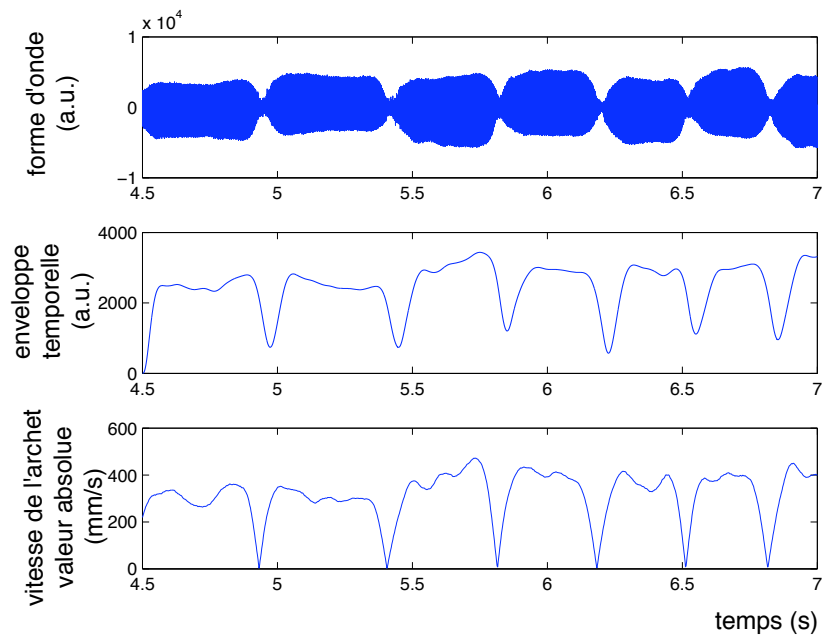


FIG. 2.5 – Corrélation entre vitesse et enveloppe du signal audio. (*haut*) : Forme d'onde, (*milieu*) : enveloppe sonore, passe-bas à 15Hz, (*bas*) : valeur absolue de la vitesse de l'archet par rapport au violon.

approche est attractive car elle permet idéalement de s'affranchir de systèmes de captation parfois encombrants pour mesurer les gestes instrumentaux. Toutefois, la difficulté se trouve déplacée dans l'utilisation et la robustesse des méthodes utilisées pour traiter le signal acoustique : par exemple, le calcul de l'enveloppe sonore d'un signal pose de manière classique le problème de la taille de la fenêtre de lissage à utiliser. Néanmoins, en partant du signal acoustique pour aller vers les gestes, cette approche offre un angle d'attaque intéressant pour étudier les liens entre les gestes et le son.

Comparaison entre capture du mouvement et capteurs embarqués.

Concernant la précision des mesures et la gêne pour les instrumentistes, les systèmes de capture de mouvement (e.g. *VICON*) sont nettement supérieurs. Ils permettent notamment de mesurer la position de plusieurs marqueurs au cours du temps autorisant l'enregistrement simultané des mouvements du corps de l'instrumentiste et de l'archet, ou de la flexion de la baguette par exemple. Par différentiations successives, il est ainsi possible de calculer vitesses et accélérations des marqueurs. Toutefois, le processus de dérivation numérique diminue le rapport signal / bruit et il est souvent nécessaire d'avoir recours à un filtrage passe-bas⁵ pouvant potentiellement affecter la précision des mesures. Un tel cas est illustré sur l'accélération de l'archet d'un violoniste sur la figure 2.6. Deux jeux de paramètres sont utilisés pour calculer l'accélération de l'archet à partir de la mesure de la position par un système *VICON* à 500Hz à l'aide d'un filtrage de Savitsky-Golay : le premier utilise un ordre d'interpolation trop faible et le deuxième trop élevé. Les deux signaux obtenus peuvent être comparés au signal enregistré par un accéléromètre. Dans le premier cas, l'accélération obtenue est lisse mais les pics d'accélérations sont temporellement plus étalés et moins prononcés. Dans le deuxième cas, les pics sont plus précis, mais des pics supplémentaires apparaissent en artefact. Il convient alors de choisir attentivement les va-

⁵la différentiation et le filtrage peuvent être effectués simultanément à l'aide d'un filtrage Savitsky-Golay, effectuant la différentiation d'une interpolation polynomiale du signal d'origine.

leurs des paramètres (taille de fenêtre et ordre de modélisation) pour réaliser le meilleur compromis entre précision et artefacts : ce choix dépendant du contenu fréquentiel de l'accélération doit être adapté à chaque cas. Cet aspect représente la limitation majeure des systèmes de capture de mouvement. Lorsque le filtrage est correctement adapté, les mesures effectuées avec le système *VICON* et les accéléromètres sont remarquablement proches, comme le montre la Figure 2.6 (bas) où l'erreur est inférieure à 3%. Néanmoins, les modules de capteurs embarqués de l'IRCAM donnent des mesures légèrement plus précises d'accélération. De plus, leur souplesse d'utilisation permet de les utiliser dans différents lieux ce qui facilite les campagnes de mesures. Cependant, le calcul de l'accélération à partir des signaux des accéléromètres peut être compliqué en particulier lorsque les variations d'angle de l'archet sont conséquentes (e.g. changement de cordes).

Les mesures utilisées dans les articles I et II ont été effectuées grâce aux modules sans fil de l'IRCAM en partie au Conservatoire National de Région à Dijon, où étaient les élèves, et en partie à l'IRCAM. Elles se concentrent sur les accélérations de l'archet sur chaque corde individuellement, permettant ainsi de considérer les accélérations relatives sur chaque coup et entre les coups d'archet. Les mesures utilisées dans les articles III et IV ont été effectuées simultanément avec un système *VICON*, dans le cadre de collaborations avec l'IRISA Rennes et l'IDMIL de l'Université de McGill, Montréal, et un module sans fil. Elles offrent un jeu de mesures précises des mouvements de l'archet et du corps des instrumentistes.

2.1.3 Visualisations des données

Les différents systèmes de captations peuvent fournir des mesures très précises du mouvement. Toutefois, les données issues des systèmes de captation tel que *VICON* peuvent être très nombreuses, compliquant ainsi leur analyse. Dans le cas du violon, neuf marqueurs sont placés sur l'instrument (violon et archet), et chaque marqueur a quatre dimensions (espace, temps). De plus, le nombre de marqueurs augmente rapidement dès qu'il s'agit de

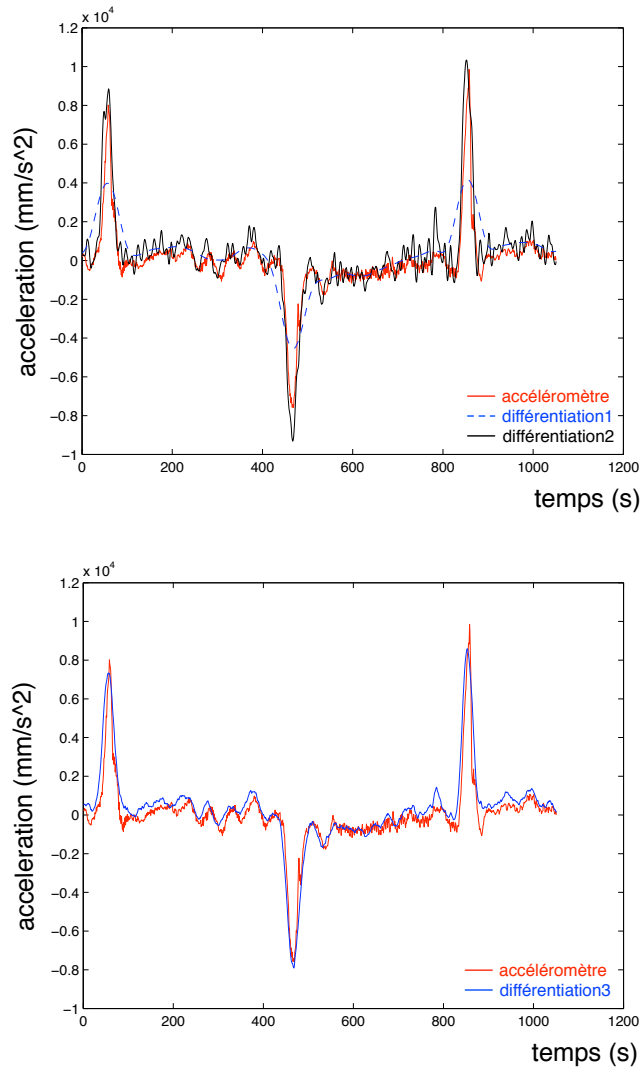


FIG. 2.6 – Comparatif entre le signal d'un accéléromètre et des accélérations obtenues à partir de trois différentiations numériques différentes d'une mesure de position effectuée par un système *VICON*. *différentiation1* : taille de fenêtre=101 éch., ordre=2. *différentiation2* : taille de fenêtre=41 éch., ordre=11. Meilleur compromis *différentiation3* : taille de fenêtre=41 éch., ordre=2.

capter également le corps de l'instrumentiste. Les possibilités de visualisation des données deviennent alors critiques.

Les travaux reportés dans l'article (Schoonderwaldt and Wanderley, 2007), permettent une visualisation remarquablement intuitive des données du mouvement des instrumentistes à cordes, à l'aide de projections dans divers espaces incluant la vitesse de l'archet, l'angle de l'archet avec les cordes, les déplacements de l'archet par rapport au violon dans chaque direction de l'espace, la pression de l'archet sur les cordes.

D'autres visualisations sont possibles parmi lesquelles un diagramme {vitesse de l'archet, accélération de l'archet}, espace des phases, tel que représenté figure 2.7. Cette visualisation permet de résumer les aspects dynamiques de l'archet, notamment les relations entre vitesse et accélération. Par exemple, les deux modes de jeu *Détaché* et *Martelé* sont représentés de manière distincte : à chaque changement de sens de l'archet, c'est-à-dire lorsque la vitesse change de signe, les accélérations sont maximales pour les coups en *Détaché*, et sont nulles pour *Martelé*. Ce diagramme illustre de manière claire que les coups *Martelé* commencent et terminent au repos tandis que les coups d'archet *Détaché* sont, dans cet exemple, produits à la suite.

Des alternatives intéressantes de navigation dans les données sont également proposées en s'appuyant sur une modalité sonore plutôt que visuelle, comme reportés pour les gestes ancillaires dans l'article (Verfaillie et al., 2006a).

2.1.4 Des contextes musicaux et non-musicaux : protocole

Les études menées dans cette thèse s'appuient sur la mesure et l'analyse du jeu d'instrumentistes à cordes. L'idée générale de la méthodologie est de créer des situations de jeu permettant d'examiner différents aspects de la maîtrise d'archet des musiciens et du rapport avec la vibration de la corde. La méthodologie repose en particulier sur la mesure du mouvement de l'archet pendant l'exécution de modes de jeu sous deux aspects : une forme "prototypique", avec des notes isolées et des gammes, et une forme "évoluée", prise dans des contextes musicaux plus complexes et des fragments de pièces

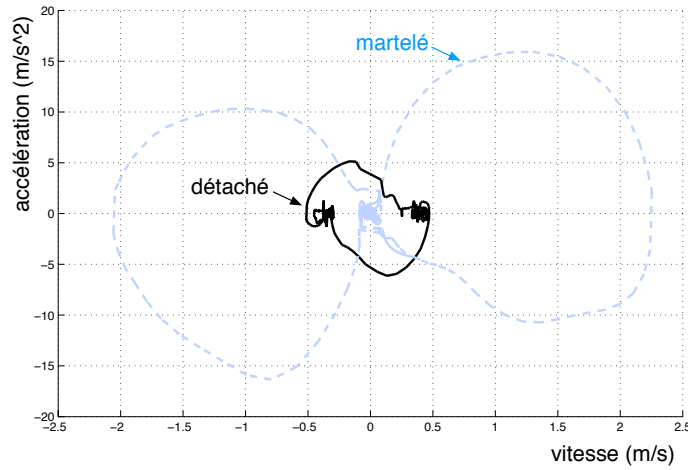


FIG. 2.7 – Diagramme de phase (vitesse, accélération) pour les modes de jeu *Détaché* et *Martelé*.

du répertoire. L'étude des modes de jeu sous leur forme prototypique permet d'établir un référent d'analyse auquel les mêmes modes de jeu sous leur forme évoluée peuvent être comparés. Ensemble, ces deux points de vue peuvent offrir un panorama large des stratégies de contrôle des instrumentistes. Ainsi, il est possible d'apprécier l'influence de paramètres tels que la nuance, le tempo, le contexte d'exécution sur la production de modes de jeu. D'une manière plus générale, l'idée directrice dans l'étude de l'influence de ces variations de paramètres sur les mouvements d'archet est de pouvoir dégager des notions d'invariants et de variances, notions importantes à la compréhension du contrôle de l'instrumentiste puisqu'elles renvoient aux questions plus fondamentales de "ce qu'on doit faire" et "ce qu'on peut faire" pour effectuer un mode de jeu.

Egalement, l'une des hypothèses supposées dans la méthodologie est qu'à partir de l'analyse de quelques modes de jeu choisis, il est possible de déduire des stratégies plus générales sous-tendant d'autres modes de jeu. Ainsi, dans cette thèse, l'étude se concentre délibérément sur un nombre restreint de modes de jeu sélectionnés et utilisés comme points cardinaux dans l'espace

des techniques d’archet.

2.2 Article I

Dans leurs travaux impliquant respectivement les mouvements de percussionnistes et de pianistes, (Dahl, 2005) et (Goebl et al., 2005) ont créé différentes situations musicales pour l’interprète en faisant varier les nuances et les modes de jeu, de manière à dégager les similitudes et différences entre interprètes (Dahl, 2005) et dans le comportement des pianos (Goebl et al., 2005).

Dans cet article, les trois modes de jeu des cordes frottées *Détaché*, *Martelé* et *Spiccato* sont ainsi étudiés dans le contexte de gammes jouées à différentes nuances et différents tempos. L’un des objectifs est d’étudier pour un instrument à son entretenu comme le violon dans quelle mesure il est possible d’extraire des paramètres du mouvement de l’archet permettant de caractériser à la fois l’invariance et la variation de chacun des modes de jeu dans ces contextes.

Construction de la base d’étude

Deux violonistes sont enregistrés pendant l’exécution de gammes dans chacun des trois modes de jeu *Détaché*, *Martelé* et *Spiccato*. Des variations d’exécution sont demandées au niveau du tempo (60bpm et 120bpm) et au niveau des nuances (*pianissimo*, *mezzo forte* et *fortissimo*). Les mesures sont effectuées à 200Hz à l’aide d’accéléromètres placés au talon de l’archet. La vitesse de l’archet est obtenue à partir des signaux d’accéléromètres.

Les enregistrements du mouvement d’archet des instrumentistes constituent un flux ininterrompu de données. De manière à étudier les particularités de chaque mode de jeu, les données sont segmentées. Dans cet article, les unités d’études choisies sont les coups d’archet. Dans les articles II et IV, d’autres unités sont considérées pour décrire les modes de jeu *Détaché* et *Martelé*. La segmentation est ici effectuée à l’aide d’une détection de pics sur le signal d’accélération. La base d’étude est alors constituée des mesures

d'accélération et de vitesse de 1320 coups d'archet tirés et poussés, exécutés dans trois modes de jeu différents, à trois nuances, deux tempos et par deux violonistes.

Détermination des paramètres de description

Les quatre descripteurs a_{max} , a_{min} , v_{max} et v_{min} (premier minimum local après v_{max}) utilisés dans cet article correspondent à une paramétrisation simple de la courbe temporelle de vitesse. Ils correspondent respectivement aux accélérations au début et à la fin du coup d'archet et aux vitesses maximale et minimale sur le coup d'archet. Ensemble, ces quatre descripteurs permettent de rendre compte du comportement dynamique de l'archet pendant l'exécution des modes de jeu.

Une analyse discriminante linéaire (*linear discriminant analysis*) indique que les trois modes de jeu *Détaché*, *Martelé* et *Spiccato* décrits par a_{max} , a_{min} , v_{max} et v_{min} peuvent être séparés à l'aide de deux dimensions dans lesquelles les paramètres a_{max} et a_{min} ont une contribution majeure. Sur la base de cette observation statistique, ces deux paramètres sont alors choisis pour décrire les mouvements de l'archet pour ces modes de jeu et offrent alors une représentation pour chaque coup dans un espace à deux dimensions. Chaque mode de jeu forme un nuage de points, lui-même composé de trois sous-nuages correspondant aux nuances *pianissimo*, *mezzo forte* et *fortissimo*.

Reconnaissance

Trois tâches de reconnaissance par K plus proches voisins ($K = 10$) permettent de mettre en rapport l'invariance et la variance des paramètres a_{max} et a_{min} en fonction notamment des modes de jeu et des nuances.

La première reconnaissance est effectuée en prenant en référence un quart de l'ensemble total d'étude, i.e. mélangeant les deux instrumentistes, les trois nuances, les deux tempos (soit 320 points), et en considérant trois classes (une par mode de jeu). La deuxième reconnaissance prend en référence les coups d'archet d'un instrumentiste et essaie de reconnaître ceux du deuxième

instrumentiste, sur les trois mêmes classes. Ces deux expériences affichent des taux de reconnaissance élevés, pour la première (96,7%, 85,8% et 89,0%) pour les modes de jeu (*Détaché*, *Martelé* et *Spiccato*), et (100,0%, 100,0% et 68,7%) pour la deuxième. La troisième tâche de reconnaissance considère neuf classes, une par nuance de mode de jeu, et prend plus de points en référence, i.e. les deux tiers de l'ensemble total (880 points).

D'une manière générale, les taux de reconnaissance sont élevés. Ce résultat révèle une forte constance des paramètres a_{max} et a_{min} pour chaque mode de jeu, en particulier sur les deux premières expériences où le nombre de points en référence est relativement faible (1/3) par rapport au nombre de points testés (2/3). Ces taux confirment l'observation graphique selon laquelle les modes de jeu *Détaché*, *Martelé* et *Spiccato* forment des nuages distincts. La troisième expérience affiche toutefois des taux plus bas pour certaines classes (*Martelé ff*, *Spiccato mf*, *Spiccato ff*) et des confusions conséquentes, notamment (*Martelé pp*, *Spiccato mf*), (*Spiccato pp*, *Détaché pp*) et (*Martelé mf*, *Martelé ff*). Ces confusions révèlent néanmoins des aspects pertinents sur les modes de jeu comme expliqué dans la discussion.

D'autre part, en étudiant spécifiquement les tempos 60bpm et 120bpm, aucune différence significative n'a été observée à l'aulne de la paramétrisation a_{max} et a_{min} . Toutefois l'influence du tempo, notamment sur les profils de vitesse et d'accélération, est étudiée plus en détail dans l'article III.

Discussion

Ces résultats mettent en avant plusieurs aspects sur le contrôle de l'archet dans les cordes frottées. Ils révèlent tout d'abord une constance dans l'exécution des modes de jeu *Détaché*, *Martelé* et *Spiccato* dans les expériences 1 et 2. Cette constance peut être mise en relation avec le contexte de gammes, où l'instrumentiste recherche justement une uniformité. Toutefois, il est intéressant de noter la similitude d'exécution entre les deux violonistes comme le montrent les taux élevés de reconnaissance de l'expérience 2, où les données des deux instrumentistes sont croisées. Dans ces deux cas, l'archet est soumis à une dynamique précise, rapportée dans cette étude par les rapports entre

le maximum et le minimum d'accélération de l'archet sur chaque coup (a_{max} et a_{min}).

Les confusions s'opérant dans l'expérience 3 traduisent une similitude de mouvement d'archet dans l'exécution des modes de jeu, par exemple entre *Martelé pp* et *Spiccato mf*. Cette similitude est cependant informative puisqu'elle concorde avec le point de vue des violonistes reconnaissant cette similarité gestuelle. À l'aide de la paramétrisation ($a_{max}; a_{min}$) il est ainsi possible de rendre compte du fait que les frontières entre modes de jeu sont floues et qu'il existe une certaine continuité entre ces modes de jeu. En effet, deux points proches dans l'espace ($a_{max}; a_{min}$) ont des caractéristiques sonores similaires et sont à l'écoute perceptivement proches.

D'autre part, l'accélération de l'archet joue un rôle prépondérant dans la mise en vibration de la corde, et conditionne en conjonction avec la pression de l'archet sur les cordes l'apparition de différents régimes *pre-Helmholtz* dans le mouvement de la corde (Guettler, 2004b). Les paramètres a_{max} et a_{min} peuvent ainsi être reliés à la création de différentes qualités de transitoires. Cet aspect est abordé dans l'article II.

Cette étude met en avant des paramètres statistiquement invariants dans la dynamique de l'archet participant à l'élaboration d'un prototype pour chaque mode de jeu. Elle explore également le domaine de variations de chaque mode de jeu à l'aune de paramètres issus de l'accélération de l'archet dans le sens de la baguette, en montrant que la délimitation entre modes de jeu peut être floue. Ces résultats posent des éléments de référence pour les études dans des situations musicalement plus complexes étudiées dans l'article IV.

2.3 Article II

Dans l'article I, les unités d'études choisies pour analyser les modes de jeu sont les coups d'archet. L'article II s'intéresse lui aux transitions entre coups d'archet et les unités d'études sont par conséquent définies centrées sur chaque changement de sens de l'archet. Dans cet article, l'objectif est

d'étudier la signifiante des transitions entre coups d'archet du point de vue sonore et gestuel dans la maîtrise des modes de jeu. Il traite notamment des relations entre la vibration de la corde et les mouvements de l'archet dans les modes de jeu *Détaché* et *Martelé*.

Protocole

Sept étudiants de troisième cycle au Conservatoire National de Région de Dijon et leur professeur ont été enregistrés exécutant des gammes en *Détaché* et *Martelé*, aux nuances *forte* et *piano*, à 80bpm, à l'aide d'accéléromètres placés au talon de l'archet, échantillonnés toutes les 3ms. La prise de son est effectuée à l'aide d'un micro fixé derrière le chevalet.

Les modes de jeu *Détaché* et *Martelé* sont choisis comme représentants de deux manières différentes mais complémentaires de mettre la corde en vibration. La technique *Détaché* implique que le mouvement de la corde est soutenu d'un coup d'archet à l'autre, tandis que le mouvement de la corde est brièvement stoppé par l'immobilité de l'archet avant le début d'un *Martelé*. Ces deux modes de jeu engendrent ainsi deux qualités différentes de transitions entre coups d'archet. Les exécutions du professeur et des élèves sont mises en perspectives les unes par rapport aux autres pour mettre en avant leurs maîtrises des transitions.

Bruit du transitoire

La transition d'un coup d'archet à un autre se traduit dans le mouvement de la corde par une phase transitoire pendant laquelle un régime périodique s'arrête et un nouveau se crée. Cette phase transitoire non périodique se traduit par un bruit large bande, filtré par la caisse du violon. C'est ce bruit résultant que les instrumentistes à cordes apprennent implicitement ou explicitement à contrôler pour produire différentes variations des modes de jeu *Détaché* et *Martelé*. En effet, différentes qualités de bruits de transitoires peuvent être obtenues en fonction de l'accélération de l'archet et de la pression de l'archet sur les cordes (Guettler, 2004b).

L'extraction du signal sonore de la contribution du mouvement transitoire non-périodique de la corde est effectuée indirectement en modélisant le signal sonore comme la somme d'un signal sinusoïdal exponentiellement amorti pour la partie périodique, et d'un signal à paramètres stochastiques pour la partie non périodique. En utilisant le formalisme proposé par les Méthodes à Haute Résolution (Badeau, 2005), les paramètres de la partie déterministe sont estimés avec un biais minimum sur de courtes fenêtres temporelles (toutes les $2ms$), adaptées au temps d'existence du régime transitoire en situation musicale (de l'ordre de $50 - 90ms$, Guettler and Askenfelt (1997)). Le signal résiduel obtenu par soustraction au signal sonore original du signal déterministe resynthétisé (OverlapAdd) contient alors la partie du signal original échappant au modèle déterministe et inclut la contribution due au mouvement transitoire non périodique de la corde.

Relations temporelles accélération-résiduel

Les techniques d'archet *Détaché* et *Martelé* réalisent des combinaisons temporelles particulières de composantes bruitées et de composantes périodiques, auxquelles Galamian réfère en parlant de "consonnes" et "voyelles", ajoutant que les instrumentistes doivent les contrôler pour articuler leur discours musical :

"La production du son sur les instruments à cordes ne concerne pas que le son continu. Il faut la compléter par l'alliance d'éléments percutants ou accentués qui donnent du caractère et un contour. Dans la musique instrumentale, le rapport entre ces éléments percutants et ceux de la ligne mélodique pure est le même que celui des consonnes et des voyelles dans la parole puis dans le chant. L'omniprésence des consonnes dans toutes les langues prouve qu'elles sont essentielles au langage. Ce principe doit être transposé dans la musique instrumentale où les sons percutants comme des consonnes sont souvent nécessaires pour donner une définition et une forme plus claires aux sons voyelles de la ligne mélodique.(...) Avec la main de l'archet, la consonne sera toute attaque qui n'est

pas sans heurt, comme le martelé, le détaché accentué, le spiccato” Galamian (1999, p. 21).

La réalisation de ces combinaisons bruit-son, c’est-à-dire la production des proportions voulues de ”consonne” et de ”voyelle” pour chaque coup, impose des contraintes liant notamment l’accélération de l’archet et la vibration de la corde. Dans le cas du *Martelé* la composante bruitée est générée par le relâchement d’une pression importante exercée par l’archet sur la corde accompagné d’une vive accélération. Dans le cas du *Détaché*, la composante bruitée est générée au moment du changement de sens de l’archet, alors que l’accélération est maximale. Le décalage temporel observé pendant les transitions entre les premiers moments respectifs de l’accélération de l’archet et de la puissance du résiduel (positif pour *Détaché* et négatif pour *Martelé*) permet de rendre compte de ces contraintes de contrôle inhérentes à la production de différentes qualités de transitions.

Le décalage temporel entre accélération de l’archet et puissance du résiduel est alors utilisé comme mesure relative de qualité entre différentes transitions. Comme le montre l’analyse de ses gammes *forte* et *piano*, le professeur démontre une constance d’exécution pour chaque mode de jeu. En comparaison, les élèves présentent une plus grande disparité dans leurs coups. Dans un exercice de gammes où l’uniformité est implicitement recherchée, cette disparité d’exécution résonne avec le fait que les élèves ont une maîtrise plus ou moins avancée et sont de fait plus ou moins capables de reproduire le même geste. En particulier, une asymétrie est observée dans les coups d’archet : alors que les tirer, respectivement les pousser, sont exécutés avec une certaine constance (similaire au professeur), les tirer et les pousser sont différents aux vues du décalage temporel. Cette asymétrie est traditionnellement considérée comme un défaut dans l’enseignement classique du violon.

Effectuer le mode de jeu *Martelé* à la nuance *piano* représente en soi une difficulté particulière : le son produit doit avoir une attaque brutale et rester globalement doux. Si le professeur produit une série de transitions très similaires pour *Martelé piano*, les élèves produisent des transitions très différentes sur le même exercice, certaines ne respectant plus la relation temporelle

précédemment établie entre accélération de l'archet et puissance du résiduel : à l'écoute la sonorité de ces transitions a perdu de son caractère explosif et est proche de celle d'un *Détaché*. Ce résultat offre une vue complémentaire au résultat trouvé dans l'article I sur la similarité entre *Martelé piano* et *Détaché mf*.

Discussions

Une description plus complète du comportement temporel des paramètres de l'archet et de la vibration de la corde pourrait être obtenue par la mesure de la pression. En particulier, il serait possible d'étudier les relations entre les amplitudes de l'accélération de l'archet et de la puissance du bruit, et d'étudier notamment leurs influences sur les relations temporelles décrites dans cet article. Accélération et pression sont cependant liées dans la production de différentes qualités de régimes *pre-Helmholtz* (Guettler, 2004b) : pour une évolution temporelle donnée de la vibration de la corde, les profils de pression et d'accélération sont alors inter-dépendants. Moyennant cette hypothèse, les résultats obtenus à partir de la considération de l'accélération uniquement apportent déjà à eux seuls des informations significatives sur la capacité de contrôle des transitions par les instrumentistes.

La production de qualités de transitions "acceptables"⁶ dans un contexte musical est soumise à des contraintes portant non seulement sur les paramètres de l'archet (Guettler, 2004b) mais également sur la durée des transitions (Guettler and Askenfelt, 1997), impliquant ainsi des relations temporelles particulières entre accélération de l'archet et bruit dans le son telles que celles reportées dans cet article pour les transitions explosives du *Martelé* et les transitions du *Détaché*. La constance des relations temporelles entre l'accélération de l'archet et le bruit dans le son reportée dans cet article indique une capacité de contrôle reproductible, conscient ou inconscient, des transitions entre coups d'archet. Les différences de variabilité entre le professeur et ses élèves peuvent être imputées à leur maîtrise respective de la corde pendant les transitions, contrôle dont l'acquisition fait partie du travail des

⁶d'après le jugement émis par 20 instrumentistes à cordes professionnels et étudiants.

musiciens (Flammer and Tordjman, 1988; Menuhin, 1973).

Ce résultat apporte un point de vue complémentaire aux résultats de l'article I, où les unités d'études correspondent à chaque coup d'archet. Ensemble, les articles I et II offrent deux segmentations temporelles des mouvements de l'archet similaires à celles opérées traditionnellement en parole séparant syllabes et diphtongues. Une segmentation temporelle s'appuyant sur la même analogie a été proposée pour les signaux musicaux dans (Schwarz, 2004), distinguant notes et "di-notes", mais reste encore très rarement utilisée. Les résultats reportés dans cet article indiquent ainsi que la transition entre les coups d'archet est aussi significative que les coups eux-mêmes.

2.4 Article III

L'article I a étudié l'influence de la nuance sur la production des modes de jeu et a notamment dégagé la notion de continuité entre modes de jeu. L'article III approfondit l'influence du tempo sur le contrôle de l'archet des instrumentistes à corde et vient élaborer cette notion de continuité. L'objectif est de tester une hypothèse élaborée à partir des connaissances empiriques de violonistes, selon laquelle il existe "différents"⁷ mouvements d'archet en fonction de la fréquence des coups d'archet. Cette hypothèse se pose en référence à des travaux sur la voix chantée concernant l'existence de différents mécanismes laryngés dépendants de la hauteur de la note (Henrich, 2006), et en contrôle moteur concernant la transition entre modes de coordinations en fonction de la fréquence d'exécution (Haken et al., 1985). L'article III propose ainsi une étude exploratoire testant cette hypothèse.

Dans cet article, les mouvements d'archet ainsi que les mouvements du haut du corps de deux violonistes⁸ et d'un altiste sont enregistrés à l'aide d'un système VICON, à 500Hz. Il leur est demandé d'exécuter un *accelerando* /

⁷les instrumentistes parlent de modification de la raideur du bras : les mouvements rapides peuvent demander à ce que le bras droit soit particulièrement "élastique".

⁸un violoniste supplémentaire a été enregistré et les analyses de ses mouvements viennent compléter les résultats présentés dans l'article III.

decelerando en partant d'un *Détaché* à tempo modéré ($1Hz$) jusqu'au plus rapide qu'ils peuvent exécuter (environ $16Hz$). La pression d'archet est mesurée à l'aide d'un capteur spécifique construit à l'Ircam (Demoucron et al., 2006) à $500Hz$. Le son des instruments est enregistré simultanément et de manière synchrone au mouvement.

Changement du mouvement d'archet

Le changement de tempo effectué par les instrumentistes se traduit tout d'abord par une variation proportionnelle de la longueur d'archet utilisée, tel qu'observé par (Winold et al., 1994). La vitesse et l'accélération sont notamment contraintes par le déclenchement et le maintien d'un régime périodique de la corde (Guettler, 2004b) et conservent une amplitude similaire tout au long de l'exercice. De plus, l'accélération imposée à l'archet, contrainte par les limites physiologiques de l'instrumentiste, est bornée (Nelson, 1983) : la combinaison tempo rapide et longs coups d'archet est effectivement difficile à réaliser.

Deux profils différents de vitesse et d'accélération apparaissent au cours de l'*accelerando* / *decelerando*. Le changement d'un profil "vitesse rectangulaire" à un profil "vitesse sinusoïdale" s'opère pendant l'*accelerando* et se fait de manière abrupte, comme le révèle un suivi de la forme du profil effectué à l'aide d'une régression non-linéaire sur un demi-sinus. Le changement inverse s'effectue pendant le *decelerando*.

Changements dans le mouvement du bras

Bien que chaque instrumentiste présente des mouvements de bras différents, un changement similaire apparaît également au niveau des articulations du bras comme en témoignent notamment les dérivées de l'angle du coude : d'un profil complexe, elles passent abruptement à un profil sinusoïdal.

De plus, au plus rapide de l'*accelerando* / *decelerando*, une inversion de phase s'opère entre l'angle du poignet et l'angle du coude pour les deux instrumentistes. Toutefois ce changement dans le mouvement du bras n'a

pas de répercussion directe sur les mouvements de l'archet, comme le montre l'accélération de l'archet. Ce dernier résultat laisse alors envisager une réorganisation du mouvement telle que celles fréquemment étudiées en contrôle moteur (Haken et al., 1985).

Influence sur le son

L'*accelerando* / *decelerando* est effectué sur une seule note (do 4). La différence trouvée en comparant le spectre de trois notes avant et trois notes après le changement de profil observé sur le mouvement de l'archet est relativement faible $1dB$ en moyenne avec un écart-type de $4dB$. Les partiels avant et après le changement ont une fréquence et une amplitude similaires et une différence est observée sur le niveau de bruit (résiduel) dans les fréquences moyennes (entre $100Hz$ et $600Hz$). Ces résultats indiquent que la différence sonore est subtile alors que le changement dans le mouvement de l'archet est conséquent.

Discussion

Les analyses effectuées sur les trois instrumentistes indiquent l'existence de différents profils de vitesse et d'accélération de l'archet lorsque le tempo varie. Pour valider la généralité des résultats, il faut bien entendu mener des analyses complémentaires sur un plus grand nombre d'instrumentistes. Néanmoins, l'étude effectuée ici soulève déjà des points intéressants à discuter. En particulier, les résultats indiquent une corrélation entre le tempo et le profil de vitesse : le profil initialement rectangulaire devient sinsusoïdal pour les tempos les plus rapides. Toutefois, l'hystérésis notée dans le changement d'un profil à l'autre indique qu'il existe un intervalle de tempo (approximativement entre $4Hz$ et $6Hz$) dans lequel les deux profils peuvent coexister, soulevant ainsi la question de ce qui est imposé par le tempo et ce qui reste sous le contrôle des instrumentistes. Les contraintes d'origine biomécanique liées au mouvement sont étudiées dans l'article IV et apportent des éléments de réponse à la première partie de la question.

Le changement observé dans les profils de vitesse et d'accélération de l'archet pour les tempos plus rapides ne s'accompagne pas d'un changement aussi tranché sur le spectre des notes produites. Des analyses sur des instrumentistes de niveaux différents permettront de déterminer si cette similitude sonore est due à l'expertise des musiciens étudiés dans cet article qui, de même que les chanteurs expérimentés peuvent gommer la discontinuité de hauteur lors d'un changement de mécanisme laryngé, peuvent minimiser l'effet sonore du changement en jouant sur d'autres paramètres (pression, position sur la corde). Il peut également être avancé que le changement de profil des mouvements de l'archet observé dans cette étude agit sur des aspects plus fins du timbre, e.g. sur la qualité des transitions entre notes, de manière suffisamment subtile pour permettre aux instrumentistes d'utiliser deux mouvements d'archet différents pour produire un son "similaire".

Le deuxième changement découvert dans les articulations du bras s'opère alors que les articulations du coude et du poignet ont un mouvement périodique sinusoïdal. Des analyses sur un grand nombre de musiciens effectuant un *accelerando* métré permettront de déterminer si le changement de phase observé correspond à une réorganisation de la coordination des articulations du bras imposée par le tempo élevé et le cas échéant d'en préciser les conditions d'apparition.

Cet article prend en compte des considérations biomécaniques sur les instrumentistes. Les résultats indiquent l'existence de différents mouvements à tempo faible et à tempo élevé définissant ainsi deux espaces de possibles séparés. En conséquence, il apparaît que les contraintes définies par l'acoustique de l'instrument et la biomécanique ne tendent pas à définir un seul espace mais plutôt plusieurs espaces qu'un instrumentiste doit apprendre à maîtriser.

2.5 Article IV

Plusieurs travaux ont dégagé des comportements d'anticipation dans les mouvements d'instrumentistes, notamment pour les pianistes (Engel et al.,

1997; Loehr and Palmer, 2007) et pour les percussionnistes (Dahl, 2000). L'article IV propose d'étudier l'existence de tels comportements dans le cas des instrumentistes à cordes, où le contrôle sur le son est continu. Pour mener cette étude, différents modèles inspirés de travaux en biomécanique sont utilisés pour comparer les mouvements effectués par des instrumentistes, dans différentes situations musicales. Les résultats de cet article permettent de placer les résultats des articles I et III sous un éclairage différent.

Présentation des modèles

Le formalisme adopté est emprunté aux travaux de (Nelson, 1983) : une masse m , en déplacement linéaire sur une distance D , pendant un temps T , soumise à une force par unité de poids $u(t)$ bornée. Le déplacement de la masse commence et termine par une vitesse nulle.

La considération de contraintes supplémentaires relatives à une optimisation de coût de performance définit différentes classes de solutions aux profils de vitesses particuliers : minimisation d'énergie, de la force, du temps, du *jerk* (dérivée de l'accélération).

Cet article utilise ce formalisme pour synthétiser des profils de vitesse d'archet correspondant à différents coûts et effectue la comparaison avec différents profils de vitesse d'archet mesurés présentant des caractéristiques particulières, dans le cas de modes de jeux *Détaché* et *Martelé* (article I), ou en fonction du tempo (article III). Des situations musicales plus complexes sont étudiées.

Mesures

Le dispositif de captation du mouvement est le même qu'utilisé dans l'article III : système Vicon mesurant les mouvements de l'archet à 500Hz. Cinq exercices sont demandés à trois interprètes. Ils mettent en oeuvre des contextes musicaux simples et plus complexes. Le premier exercice demande aux interprètes d'exécuter des notes isolées, séparées par un silence, dans les modes de jeu *Détaché* et *Martelé*. Le deuxième leur demande d'exécuter une

gamme à une octave, montante et descendante, dans les modes de jeu *Détaché* et *Martelé*. Le troisième exercice consiste à mélanger différents rythmes, i.e. noire quatre doubles. Le quatrième exercice est l'exercice de l'*accelerando* / *decelerando* de l'article III. Le dernier exercice consiste à mélanger dans une même phrase les modes de jeu *Détaché* et *Martelé*.

Définition des modèles utilisés

Trois modèles inspirés des articles (Nelson, 1983; Hogan, 1984) sont re-synthétisés à partir des données capturées et comparés aux profils mesurés (cf. Figure 3 et 4 de l'article). Le modèle trapézoïdal *Trap* est inspiré de la solution *minimum impulse* : sa forme minimise les variations de vitesse sur le coup d'archet. Les deux modèles inspirés des solutions de *minimum jerk* tiennent compte des mouvements discrets J_d et cycliques J_c (Hogan and Sternad, 2007), minimisant les variations d'accélération sur le coup d'archet. Les comparaisons de données sont effectuées à l'aide d'une corrélation entre les accélérations de l'archet modélisés et mesurées.

Coups d'archet isolés et enchaînés

Pour les trois instrumentistes et pour les deux techniques d'archet, les coups d'archet isolés sont le mieux décrits par le modèle J_d . Ce résultat coïncide avec les résultats en biomécaniques portant sur les mouvements volontaires du bras (Hogan, 1984).

En revanche, en enchaînant les coups d'archet, les mouvements du *Détaché* deviennent plus proches de *Trap*, alors que ceux de *Martelé* sont toujours plus proches de J_d . Toutefois, dans les enregistrements de gammes, il existe toujours une pause entre deux coups *Martelé*, conservant ainsi une nature discrète, contrairement aux gammes *Détaché*.

Mélange de techniques d'archet

Les résultats montrent que la meilleure modélisation de la séquence (*Détaché*, *Détaché*, *Martelé*, *Martelé*) diffère de (*Détaché*, *Détaché*, *Détaché*, *Dé-*

taché) à partir du deuxième coup d'archet : le deuxième *Détaché* est mieux expliqué par J_c dans 50% des cas. Cet aspect révèle un comportement d'anticipation dans les mouvements de l'archet du dernier *Détaché* pour effectuer le premier *Martelé*. Inversement, le premier *Martelé* est parfois mieux expliqué par J_c au lieu de J_d dans 23.5% des cas. Cet aspect suggère une persistance du mouvement du dernier *Détaché* sur le mouvement du premier *Martelé*. Ces deux aspects peuvent être rapprochés de la coarticulation entre phonèmes dans la parole.

Exécution de rythmes lents-rapides

Pour deux instrumentistes l'exécution des cellules rythmiques (noire, quatre doubles) sont le mieux décrites par la succession des modèles (*Trap*, J_c , J_c , J_c , J_c). Le fait qu'un même modèle explique les quatre doubles invite à penser qu'elles peuvent être regroupées dans un même "geste". De manière intéressante, pour le troisième instrumentiste, le modèle J_c explique l'intégralité des mouvements d'archet sur les cellules : en suivant le même raisonnement que pour les deux autres musiciens, il peut être supposé que le troisième instrumentiste a exécuté la succession de cellules rythmiques dans un même "geste".

En étudiant l'*accelerando* / *decelerando* dans ce formalisme, il est possible de retrouver les deux mouvements d'archet différents présentés dans l'article III. De plus, les délimitations coïncident de manière plutôt évidente concernant la première et de manière plus discutable pour la deuxième.

Discussions

Cet article permet de considérer les efforts associés aux mouvements d'archet et ainsi d'analyser certaines stratégies utilisées par les instrumentistes experts à des fins expressives. Les résultats d'analyse sont homogènes sur les trois instrumentistes : les différentes techniques d'archet peuvent être modélisées par différentes minimisations de coûts. De plus, différentes optimisations sont utilisées en fonction du tempo comme rapporté dans l'article

III.

Néanmoins, des différences significatives entre instrumentistes sont observées dans le cas de contextes musicaux plus complexes, mélangeant des rythmes ou des techniques d'archet. Intuitivement, de telles différences pouvaient être prévisibles. La méthodologie proposée dans cet article permet de rendre compte de telles différences, et fournit un cadre d'analyse des stratégies sous-jacentes, comme par exemple la coarticulation entre mouvements d'archet.

Plus généralement, la considération de contraintes liées à l'effort peut être utile pour la définition de contrôles sur le son, s'appuyant traditionnellement sur le couple MIDI/ADSR, en groupant plusieurs unités décrites par le même effort. Enfin, cette approche peut également être pertinente pour simplifier le contrôle de synthèse sonore : elle fournit des classes de profils temporels ayant un sens physique et pouvant être contrôlés intuitivement à l'aide d'un nombre restreint de paramètres.

Chapitre 3

Discussion générale et applications

3.1 Discussion générale

Les travaux effectués dans cette thèse étudient le contrôle de l’archet des instrumentistes à cordes frottées en situations réelles de jeu sous des éclairages statistiques, acoustiques et biomécaniques, en s’appuyant sur des traités de techniques du violon écrits par des musiciens. Les résultats présentés dans les articles amènent à discuter différents aspects. Cette section met en regard les différents articles en regroupant ces aspects sous différents thèmes parmi lesquels le contrôle des dynamiques de l’archet, les différentes segmentations temporelles des données gestuelles, l’importance du tempo d’exécution et plus généralement des contraintes temporelles. Ces aspects invitent à discuter les implications des résultats notamment pour la modélisation des enveloppes sonores et à mener une réflexion plus générale sur la conception de nouvelles interfaces. Les applications réalisées pour la création musicale et potentielles pour la pédagogie sont finalement présentées.

Sur le contrôle des dynamiques de l’archet

Les résultats de l’article I montrent que statistiquement les deux instrumentistes considérés sont capables de définir différentes dynamiques d’ar-

chet pour réaliser les modes de jeu *Détaché*, *Martelé* et *Spiccato*, comme en témoignent les tâches de reconnaissance effectuées à partir des deux extrema d'accélération sur chaque coup ($a_{max}; a_{min}$). Du point de vue de la dynamique du mouvement, les deux extrema d'accélération ($a_{max}; a_{min}$) correspondent au contrôle donné par l'instrumentiste pour lancer l'archet d'une part et pour le stopper d'autre part, tel que décrit dans l'article IV. Ce contrôle sur l'archet est particulièrement important dans la maîtrise de l'archet comme le note Mari Kimura¹ (Mari Kimura, conversation privée), pour qui les élèves violonistes devraient prêter plus attention à la manière dont ils arrêtent leur archet. D'une manière plus générale, l'espace engendré par ($a_{max}; a_{min}$) offre alors un espace de représentation des dynamiques de l'archet, pour chaque coup donné, dans lequel les modes de jeu *Détaché*, *Martelé* et *Spiccato* s'expriment en différentes locations.

Les taux obtenus lors des tâches de reconnaissance indiquent également des confusions (expérience 3). Ces confusions statistiques correspondent à des points proches dans l'espace ($a_{max}; a_{min}$) mais exécutés en tant que modes de jeu différents (*Martelé pp* ou *Détaché mf*). Dans l'espace ($a_{max}; a_{min}$), la proximité entre deux points s'interprète par le fait que les deux coups d'archet correspondants ont des extrema d'accélération d'archet proches. Une distance faible dans cet espace correspond alors à des profils d'accélération ayant des dynamiques proches et peut ainsi être mise en regard avec une similarité de contrôle de l'archet. Ce point peut apporter une explication au point de vue des instrumentistes jugeant les confusions obtenues pertinentes qui pour les coups d'archet confondus ont fourni des accélérations du même ordre.

Il est ainsi intéressant de remarquer deux niveaux dans les résultats. D'une part, les deux instrumentistes de l'étude ont des domaines définis pour chaque mode de jeu à une nuance donnée et d'autre part, il existe une certaine "continuité" entre les domaines, mise en avant dans l'étude de manière indirecte par l'exécution de différentes nuances. Il est alors possible de définir

¹Mari Kimura est compositrice, violoniste et professeur à la *Julliard School of Music*. Nous avons eu de nombreux échanges sur différents aspects (techniques, expressifs) touchant à la maîtrise du violon. Une collaboration est engagée sur la création d'une pièce pour *violon augmenté*.

deux points de vue sur l'espace $(a_{max}; a_{min})$. Le premier consiste à considérer des classes définies à partir de domaines de variations des deux paramètres (e.g. les modes de jeu *Détaché*, *Martelé* et *Spiccato*) et de permettre ainsi l'opération de reconnaissances de formes sur les mouvements de l'archet : l'espace $(a_{max}; a_{min})$ devient alors discret. Le deuxième consiste à considérer directement les paramètres $(a_{max}; a_{min})$ permettant ainsi d'obtenir une description continue des dynamiques de l'archet coup par coup.

Inversement, il est intéressant de se demander si les instrumentistes sont capables de naviguer consciemment dans l'espace $(a_{max}; a_{min})$. Les paramètres $(a_{max}; a_{min})$ sont des paramètres intuitifs du contrôle d'archet. Il est alors possible de supposer² que les instrumentistes peuvent arriver à se déplacer consciemment dans l'espace $(a_{max}; a_{min})$ en développant des stratégies adaptées plus nécessairement reliées à des modes de jeu existants.

De manière à étudier la généralisation éventuelle des résultats à un plus grand nombre d'instrumentistes, l'étude menée dans l'article I sur deux instrumentistes a été reproduite sur douze autres instrumentistes issus d'une classe de troisième cycle au Conservatoire National de Région de Dijon. Si d'une manière générale les rapports entre a_{max} et a_{min} obtenus dans l'article sont conservés pour *Détaché* et pour *Martelé*, des différences idiosyncratiques apparaissent, en particulier sur le mode de jeu *Spiccato*, comme illustré sur la figure 3.1. A la différence des deux premiers, ce dernier est un mode de jeu où l'archet ne reste pas en contact permanent avec la corde mais où il la frotte uniquement une partie du temps du coup d'archet. Ainsi, pendant l'exécution et en fonction des élèves, l'archet est parfois au dessus de la corde parfois sur la corde au moment du changement de sens, générant ainsi une plus grande variabilité d'exécution. Cette différence inter-sujets peut être rapprochée des résultats obtenus dans (Dahl, 2000) où des différences entre percussionnistes sont rapportées : en alternant contact avec la corde et mouvement libre dans l'air, le mouvement de l'archet en *Spiccato* peut alors dans une certaine mesure être comparé à celui de la baguette venant frapper la peau. Toutefois,

²en tant que violoniste, je me suis plusieurs fois amusé à me déplacer dans l'espace $(a_{max}; a_{min})$ en essayant d'impulser le bon mouvement à l'archet pour atteindre un point donné, ce à quoi j'arrivais en général après quelques tâtonnements.

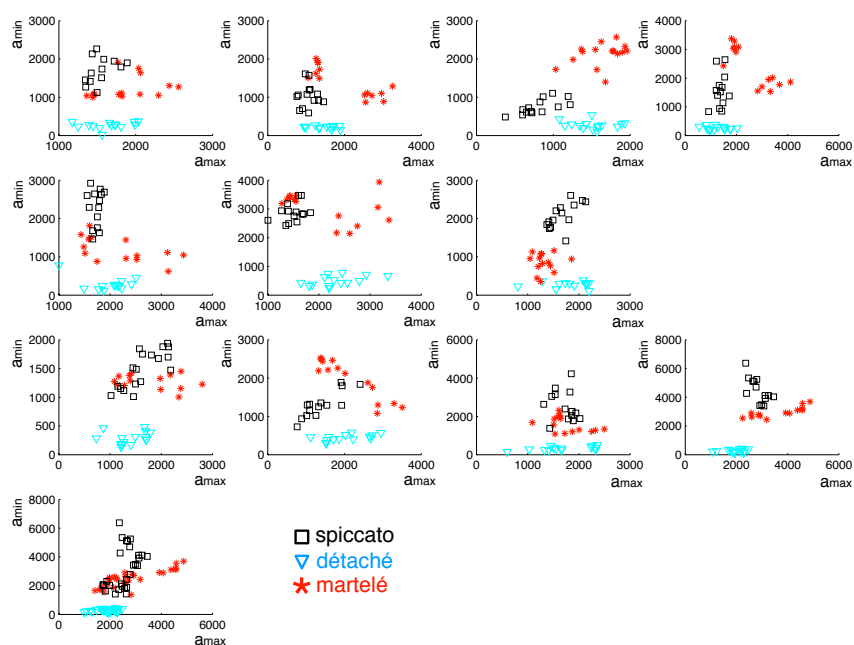


FIG. 3.1 – Espace de représentation $(a_{max}; a_{min})$ pour une classe d'élèves. Nuance *forte* et tempo à 80bpm.

indépendamment de cette dernière observation, il est intéressant de noter que chaque instrumentiste démontre une organisation plus ou moins cohérente des trois modes de jeu où chacun, y compris le *Spiccato*, se distingue des autres.

Cependant, pour les modes de jeu *Détaché* et *Martelé* où l’archet reste majoritairement en contact avec la corde, des similarités inter-sujet peuvent être observées : les rapports entre a_{max} et a_{min} sont comparables d’un sujet à l’autre. Il est possible que le fait que les instrumentistes appartiennent à une même classe de violon contribue à cette similarité. Toutefois, il est également possible de supposer l’hypothèse suivante. Pour les modes de jeu *Détaché* et *Martelé*, les instrumentistes doivent déclencher la mise en vibration de la corde, l’entretenir et la stopper en restant à la corde, contrairement au *Spiccato*. Il est alors raisonnable de penser que cette contrainte supplémentaire conditionne d’avantage les variations possibles de l’accélération de l’archet. Ainsi, les contraintes acoustiques imposées par les coups d’archet à la corde, jointes aux contraintes de tempo et de nuance peuvent être à l’origine d’une certaine similitude entre les sujets.

En offrant un résumé synthétique du mouvement de chaque coup d’archet, l’espace $(a_{max}; a_{min})$ permet d’apprécier le mouvement impulsé à l’archet pendant une performance de l’instrumentiste, et reflètent les différences de dynamiques du mouvement effectuées par l’instrumentiste, desquelles découlent les modes de jeu. Les paramètres $(a_{max}; a_{min})$ offrent ainsi un miroir permettant par exemple de s’interroger sur l’uniformité et/ou la variabilité d’exécution dans le contrôle de l’archet, problématique intéressante dans un contexte pédagogique par exemple. Cette paramétrisation a notamment été utilisée dans la création musicale *BogenLied* pour violon augmenté et électronique, composée par Florence Baschet. Les détails de cette utilisation sont développés dans la section dédiée.

Sur la segmentation

Le contrôle de la vibration de la corde, i.e. de sa mise en vibration à son arrêt en passant par les évolutions multiples dans son maintien, est la préoccupation principale des instrumentistes à cordes en situation de jeu et

conditionne les mouvements de l'archet possibles (Mari Kimura, conversation privée). Pendant un coup d'archet, l'évolution de la vibration de la corde peut être séparé en deux parties : le régime stationnaire pendant lequel la corde a un mouvement périodique et un régime transitoire pendant lequel la corde passe du repos à un mouvement stationnaire ou passe d'un mouvement stationnaire à un autre. Si la partie stationnaire permet de donner lieu à la mélodie, la partie transitoire est la partie permettant d'articuler cette mélodie et de lui donner du relief par l'ajout d'éléments percutants : ainsi, de même qu'un orateur joue sur les consonnes pour donner du corps à son élocution, un instrumentiste à corde doit savoir équilibrer "consonnes" et "voyelles" pour faire ressortir son propos musical (Galamian, 1999). Dans son ouvrage *l'Art de jouer du Violon*, Menuhin distingue aussi deux phases dans le maniement d'archet correspondant aux deux parties de vibration de la corde évoquées.

Le maniement de l'archet met en jeu deux types d'action. L'un, que j'appellerai la "gifle-rebond", est un coup vivement donné dans une direction, suivi du retour également vif dans la direction opposée, à la manière d'une balle projetée contre un mur. L'autre est un "effort continu", prolongé dans une direction unique, à la manière d'un lourd fardeau qu'on pousse ou qu'on tire le long d'une colline. Le premier type d'action est à la base du changement de direction d'archet, poussé suivi de tiré ou inversement. Le second intervient entre les changements de direction ; il est à la base du coup d'archet soutenu. (Menuhin, 1973, p. 46)

Menuhin réfère à deux actions différentes mises à contribution dans le contrôle de l'archet, la "gifle-rebond" et "l'effort continu" concernant respectivement le contrôle du régime transitoire et le contrôle du régime stationnaire de vibration de la corde, mettant ainsi les deux actions sur un pied d'égalité. Par ailleurs, Flammer partage le même point de vue dans son ouvrage *le Violon* où il explique que le retour d'archet est une technique d'archet en soi (Flammer and Tordjman, 1988). Parlant de la combinaison dans le temps entre les deux actions qu'il a définies, Menuhin ajoute :

...nous tirerons l'archet avec une gifle-rebond, puis pousserons

avec une gifle-rebond, puis tirerons de nouveau avec une gifle-rebond et ainsi de suite. (Menuhin, 1973, p. 52)

De manière intéressante, il décrit dans ce passage une succession de sons par une succession de "gifle-rebonds", autrement dit, comme une succession de transitions entre coups d'archet. Ce point de vue est finalement original dans le domaine de l'informatique musicale, où une succession de notes est traditionnellement décrite par une série d'évènements correspondant à des notes ("note-on", "note-off" dans MIDI par exemple) et dans laquelle la transition entre évènements est gérée de manière implicite. C'est dans ce contexte que s'inscrit l'étude menée dans l'article II, dont l'objectif est d'étudier le contrôle des transitions entre coups d'archet. Les enregistrements des instrumentistes sont alors étudiés sous la forme de successions de transitions. La relation entre le mouvement de l'archet et la vibration de la corde pendant les transitions du *Martelé* et du *Détaché* est précise et caractéristique pour les deux modes de jeu, comme le montrent les résultats de l'article obtenus à partir du décalage temporel entre l'accélération de l'archet et la puissance du bruit dans le signal sonore. Dans les enregistrements effectués, les instrumentistes ont montré capacité à fournir une certaine constance dans le contrôle de la vibration de la corde pendant la transition d'un coup d'archet à l'autre, pour réaliser différentes proportions entre parties bruitées et parties soutenues, entre "consonnes" et "voyelles" (Galamian, 1999). D'autre part, à partir du même critère, il est possible de rendre compte d'une similarité des transitions, sur la base des relations entre geste et son, telles qu'entre *Martelé* à la nuance *piano* et *Détaché* à la nuance *forte*, rapprochement par ailleurs obtenu dans l'article I à partir d'une description pour chaque coup d'archet (a_{max} ; a_{min}).

Au travers des articles I et II, deux types de segmentations du mouvement de l'archet des instrumentistes à cordes sont présentés. Dans l'article I, les unités d'études correspondent aux coups d'archet alors que dans l'article II, les unités d'études sont centrées autour des transitions entre coups d'archet. Ces deux segmentations permettent de tenir compte à la fois de la dynamique de l'archet sur chaque coup et de la dynamique reliant les coups,

et offrent deux points de vue complémentaires du mouvement de l'instrumentiste. Cette complémentarité peut être comparée à celle reliant diphtongues et syllabes en parole, où ces deux segmentations temporelles permettent de mettre en avant pour l'une les unités formelles constitutives de la parole et pour l'autre la transition entre unités. En particulier les diphtongues tiennent compte des liaisons entre les sonorités effectuées pendant la production de parole dues au phénomène de *coarticulation*, c'est-à-dire des influences de persistance et d'anticipation dans l'appareil phonatoire des phonèmes entourant un phonème. Il est intéressant de considérer une telle analogie dans le contexte du mouvement de l'archet des instrumentistes à cordes, des phénomènes de persistance et d'anticipation étant également présents dans la production du mouvement (Loehr and Palmer, 2007; Godøy, 2004). Ainsi, le contrôle des modes de jeu *Détaché* et *Martelé* peut être envisagé sous deux angles : la manière dont l'instrumentiste commence puis termine chaque coup d'archet, ou bien la manière dont l'instrumentiste lie ou articule les coups entre eux, aspect tout aussi essentiel à la bonne production des modes de jeu. Par leur découpage respectif, les deux segmentations des articles I et II permettent notamment de considérer ces aspects.

Les contraintes biomécaniques mises en avant dans les articles III et IV peuvent également être utilisées pour définir des segments d'analyse, composés de plusieurs coups d'archet ayant une même cohérence, comme par exemple les trois phases identifiées dans l'*accelerando* ou les fragments rythmiques dans l'article IV. Il est alors intéressant de mettre en regard la réunion de plusieurs notes en une seule unité cohérente du point de vue du mouvement avec à la notion de "geste musical", au sens musicologique, réunissant un ensemble de notes en une même unité musicalement cohérente et sémantique.

Sur le temps

Le temps est un facteur important et sous-jacent dans tous les résultats présentés dans cette thèse. Dans les articles I et II, le temps apparaît d'abord à travers le tempo fixe et régulier. Il est possible de supposer que cette régularité de tempo dans l'exécution de gammes ait facilité la constance

des mouvements effectués, les interprètes étant alors dans des conditions favorables à la répétition des mêmes mouvements pour chaque mode de jeu. D'autre part, les similarités entre modes de jeu élaborées à partir de l'espace de paramètres $(a_{max}; a_{min})$ sont valides pour des coups effectués dans un tempo moyen (60bpm et 120bpm dans les études) mais deviennent discutables lorsque le tempo augmente, les profils de vitesse et d'accélération de l'archet changeant alors, comme étudié dans les articles III et IV. Dans l'article II, le temps intervient implicitement dans la contrainte de produire différentes proportions entre partie bruitée et partie stationnaire donnant ainsi différents caractères aux transitions entre coups d'archet. Dans l'article IV, il devient un paramètre explicite : pour chaque coup d'archet, différents profils de vitesse d'archet réalistes³ sont modélisés à partir du temps mis, de la longueur d'archet utilisée et des accélérations initiale et finale sur le coup, en utilisant les contraintes physiques liant ces paramètres. L'article III étudie plus spécifiquement les effets de changements de tempo sur les mouvements de l'archet des instrumentistes. Les différents profils de vitesse et d'accélération reportés en fonction du tempo dans l'article III apportent de plus un point de vue original en informatique musicale où une telle rupture de comportement n'est généralement pas exploitée dans les interfaces musicales.

Ainsi le temps impose des contraintes sur les mouvements de l'archet à différentes échelles temporelles. Il agit notamment sur les contraintes acoustiques conditionnant la mise en vibration de la corde, la qualité des transitions et la proportion entre transitions et parties soutenues. Il agit également sur les contraintes biomécaniques qui à une échelle plus large conditionnent les mouvements de l'archet pendant l'exécution de coups plus rapides comme des rythmes par exemple. Les espaces de possibles pour le contrôle de l'archet définis par les contraintes acoustiques et biomécaniques sont alors conditionnés par le temps.

³il existe une infinité de profils possibles, dont seul un sous ensemble est réalisable par un instrumentiste.

Relations sonores

Une grande partie des travaux en informatique musicale, notamment dans le domaine du traitement du signal audio, traite les transitions entre sons successifs dans un mode allant rarement au-delà du traitement binaire, consistant à détecter si la fenêtre temporelle en cours d'analyse correspond ou non à un transitoire. Ce tri des fenêtres d'analyse permet notamment de cibler les traitements ultérieurs sur les fenêtres correspondant aux parties périodiques (e.g. changement de *pitch*, de durée). Toutefois, il est intéressant de traiter également les différentes qualités de transitions entre les notes jouées par un instrument, partie contenant une part non négligeable de l'expressivité des instruments (Traube and D'Alessandro, 2005; Penttinen, 2006) et notamment à cordes frottées (article II). Par exemple, l'un des succès du moteur de synthèse développé par la compagnie Synful⁴ n'est pas tant la concaténation d'échantillons de notes jouées par des instruments que le fait de traiter explicitement la manière d'effectuer les transitions entre notes (mode *expression*). Dans le cas des cordes frottées, les composantes sinusoïdales montrent des comportements spécifiques pendant les transitoires, notamment les deux premiers partiels (Guettler, 2002a, Introduction and comments to Paper I). Dans cette thèse, les analyses se sont cependant concentrées sur les composantes bruitées du transitoire, jusque lors peu étudiées bien que contribuant substantiellement à la signature sonore des sons de cordes frottées (Guettler, 2002a). L'article II contribue dans ce sens en plaçant le bruit du transitoire dans le son du violon au centre de son étude pour apprécier différentes qualités de transitions. Il montre ainsi que les composantes bruitées présentent également des comportements spécifiques, notamment respectivement à l'accélération de l'archet.

Les résultats des articles III et IV font apparaître différents profils de vitesse d'archet. Comme présenté dans le chapitre 2 (sous-section "Captation indirecte des gestes instrumentaux"), la vitesse de l'archet contrôle la vitesse de déplacement de la corde pendant la phase d'adhérence du cycle adhérence-glissement des cordes frottées, et en conséquence est directement corrélée à

⁴*Synful Orchestra* : <http://www.synful.com>

l'amplitude du signal sonore. La question se pose alors de savoir si les résultats obtenus dans les articles III et IV concernant les influences sur les profils de vitesse d'archet du tempo et des profils voisins s'appliquent à l'enveloppe temporelle du signal sonore. Dans les cas où les perturbations introduites par la main gauche, typiquement en effectuant un *vibrato*, sont négligeables, les évolutions de l'enveloppe sonore et de la vitesse de l'archet sont très proches et en corollaire les résultats obtenus peuvent être transposés, comme illustré graphiquement sur la figure 3.2.

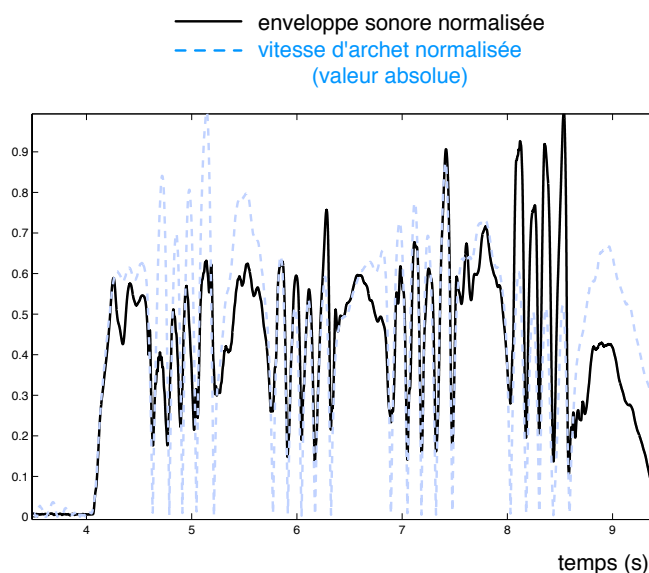


FIG. 3.2 – Enveloppe sonore et vitesse de l'archet dans le cas des rythmes présentés dans l'article IV.

Ce point de vue est original en informatique musicale où traditionnellement les modèles d'enveloppes sonores ADSR sont définies de manière indépendante et sans interaction possible entre enveloppes. D'une manière plus générale, de tels résultats sur les enveloppes sonores sont relativement difficiles à obtenir à partir de l'analyse d'un signal sonore : les enveloppes calculées à partir du signal audio combinent naturellement des informations provenant à la fois de la hauteur de la note et de l'amplitude du son (l'exemple le plus représentatif étant une fois de plus le cas du *vibrato*). L'étude de la vi-

tesse de l’archet permet ainsi de découpler les deux et d’établir des propriétés sur les enveloppes sonores.

Réflexions pour la conception de nouvelles interfaces

Les interactions entre les musiciens et leurs instruments sont complexes comme en atteste la nature multiple des contrôles présentée dans les articles : parfois continue, entre modes de jeu dans l’article I, parfois présentant des ruptures, comme dans l’article III. Les liens entre contrôle gestuel et résultat sonore sont également non-univoques, i.e. plusieurs mouvements d’archet pour une même caractéristique sonore, et évolutifs, e.g. en fonction du tempo, comme illustrés dans l’article III. Ces résultats apportent des éléments de réflexions à la définition de correspondances entre contrôleurs gestuels et sons de synthèse (*mapping*), généralement conçues de manière statique.

Les résultats des articles présentés peuvent également être utiles à la définition de contrôles de paramètres sonores réalistes⁵, notamment concernant les contraintes sur les enveloppes sonores (articles IV et III). En particulier, les notions de transitions entre notes sont relativement peu exploitées alors qu’elles jouent un rôle expressif reconnu, notamment pour les instruments à son entretenu comme les cordes frottées. Par exemple, en offrant un cadre dans lequel les notes sont représentées par un début et une fin, sans possibilité d’agir sur les qualités de transitions entre notes, la norme MIDI se retrouve particulièrement inadaptée au contrôle de sons continus comme les cordes frottées. A partir des études présentées dans cette thèse sur les cordes frottées, il est possible d’imaginer une modélisation du contrôle pour sons continus en s’appuyant par exemple sur un contrôle en vitesse pour les parties soutenues et sur un contrôle en accélération pour les transitions entre parties soutenues.

⁵Bien sûr, un tel soucis est tout à fait discutable dans le contexte artistique, le réalisme n’étant pas nécessairement le premier critère recherché. Toutefois, sa connaissance permet à la fois de le respecter et d’aller à son encontre

3.2 Applications

Cette section présente les application musicales et pédagogiques réalisées et possibles en s'appuyant sur les travaux présentés. Les travaux reportés dans l'article I ont été appliqués à la création de *BogenLied*, oeuvre de Florence Baschet pour violon augmenté et électronique. Les possibilités d'utilisation des systèmes de captation et d'analyse des mouvements sont étudiées dans le cadre du projet européen I-MAESTRO⁶.

Violon Augmenté : applications musicales

Le projet *violon augmenté* a été élaboré à partir de l'intérêt que plusieurs compositeurs ont montré par rapport au potentiel de composition que peut offrir l'utilisation d'un violon (et par la suite tous les instruments du quatuor) associé à un système de captation du geste dans des créations de musiques électroacoustiques. Un groupe de travail réunissant chercheurs, compositeurs, interprètes et réalisateurs en informatique musicale a donc été initié à l'IR-CAM créant ainsi un contexte propice à des échanges entre scientifiques et artistes autour d'un tel instrument. Parmi les créations réalisées dans le contexte de ce groupe, on trouve des compositions de Florence Baschet, Franck Bedrossian, Jérôme Combier, Juan José Eslava, Jose Miguel Fernandez, Philippe Manoury, Hèctor Parra et Sebastian Rivas. Les interprètes ayant participé et/ou participant sont notamment Jeanne-Marie Conquer (Ensemble InterContemporain), Hae-Sun Kang (Ensemble InterContemporain), Anner Mercier (Ensemble l'Itinéraire), Christophe Desjardins (Ensemble InterContemporain) et Pierre Strauch (Ensemble InterContemporain), le quatuor Danel.

Florence Baschet a pris une part très active au sein du groupe et a été notamment à l'origine de synergies très riches entre la recherche scientifique et son travail de composition. Son oeuvre pour violon augmenté et électronique, *BogenLied*, utilise directement les résultats de recherche présentés dans l'article I. Dans cette pièce, les modes de jeu *Détaché*, *Martelé* et *Spiccato* sont

⁶<http://www.i-maestro.net>

reconnus et sont utilisés pour sélectionner la transformation électroacoustique du son du violon en temps-réel. Un deuxième degré de transformation est ensuite opéré à partir du pourcentage d'appartenance à la classe du mode de jeu reconnu permettant ainsi de graduer l'effet de la transformation. La figure 3.3 illustre le principe général du contrôle des transformations sonores.

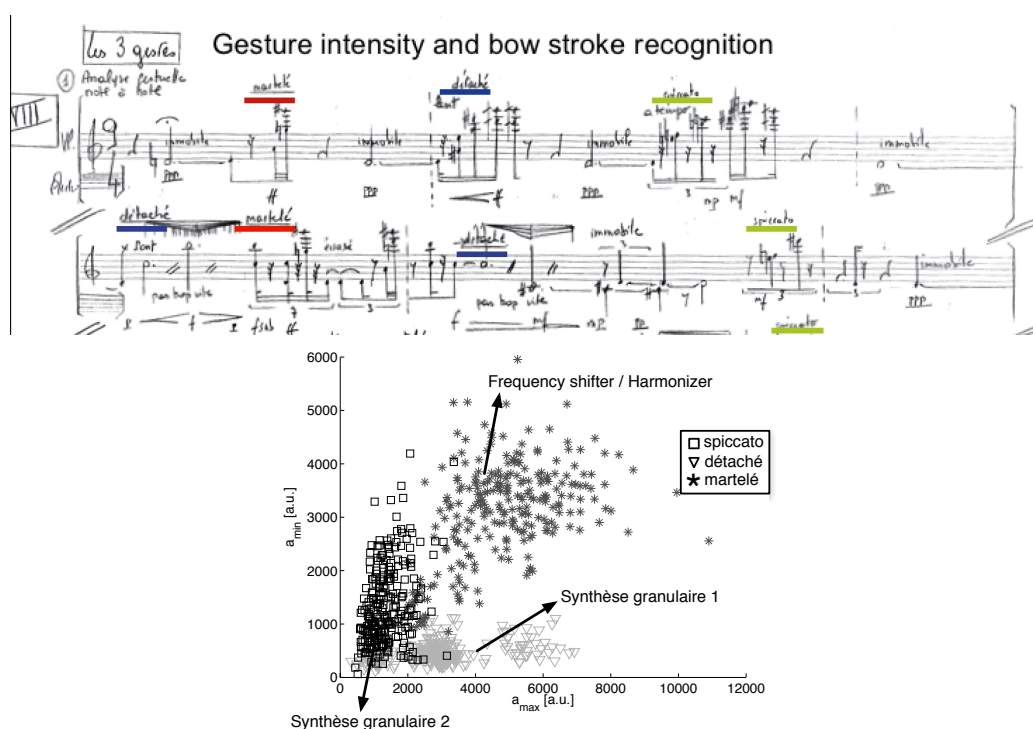


FIG. 3.3 – Les transformations sonores sont déterminées en fonction des coups d’archet de la violoniste. En haut, extrait de la partition de *BogenLied*, en bas, espace $(a_{max}; a_{min})$ et transformations sonores associées.

Cette pièce a requis l'implémentation en temps réel de l'algorithme d'analyse des modes jeu présenté dans l'article I. La figure 3.4 donne les détails de cette implémentation. Elle se constitue d'une étape pré-traitement (centrage du signal d'accéléromètre sur une fenêtre glissante, filtrage), segmentation en coups d'archet à l'aide d'une détection de pics, extraction des valeurs maximales et minimales d'accélération, et enfin reconnaissance à l'aide d'un

algorithme de K plus proches voisins.

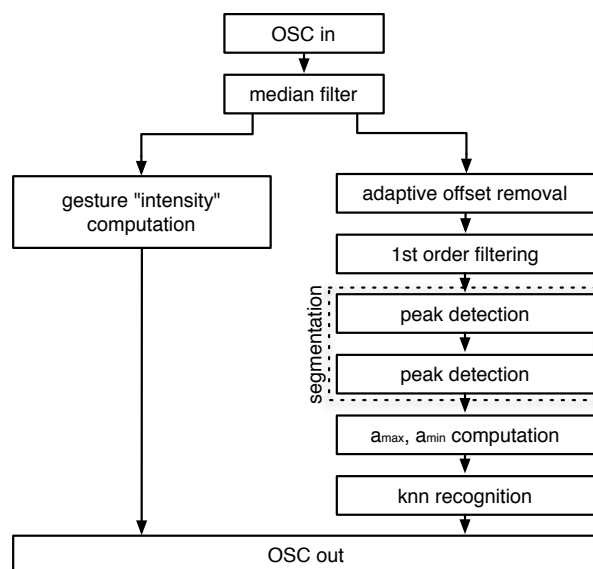


FIG. 3.4 – Diagramme explicatif de l’implémentation ”temps-réel” de l’analyse des modes de jeu.

Ainsi, le soliste peut décider et modéliser en temps réel la réponse électroacoustique. Le lien avec sa technique instrumentale lui offre un contrôle intuitif et pertinent. De plus, la soliste peut tirer partie de la finesse de son jeu pour graduer la force des effets sonores et se trouve alors en mesure d’interpréter, à chaque exécution, la pièce de manière différente au niveau acoustique et électronique. Voici la présentation de *BogenLied* que la compositrice a faite pour une émission radio du GRM diffusée sur Internet :

” *Bogen* en allemand c’est l’archet et *Lied* c’est le chant donc *BogenLied* c’est le chant de l’archet. Dans notre culture musicale on a toujours écrit les notation des instruments à cordes avec des indications d’archet : on écrit détaché, on écrit sul ponticello, on écrit martelé. Le geste de l’archet est donc un paramètre compositionnel important. (...) [Dans cette pièce] c’est l’archet qui va piloter toute l’électroacoustique. Pour l’instrumentiste c’est fabuleux, Anne Mercier le dit elle-même, elle sent qu’elle contrôle ce qui se

passé, qu'elle contrôle les transformations électroacoustiques par rapport à son geste, c'est quelque chose qui lui vient tout-à-fait naturellement. Et du coup ça permet de créer une réelle interactivité (...) ça crée une vraie relation entre elle et l'ordinateur, ça pourrait être cette relation en musique de chambre entre deux instrumentistes mais là c'est entre le soliste et l'ordinateur." (Interview pour les Ondes, GRM radio, Live electronics 07/01/06).

BogenLied a été créée par Anne Mercier lors du festival *Why Note* en novembre 2005 à Dijon et a été reprise par Anne Mercier à la Maison de la Radio pendant le festival *Live Electronics* du GRM en janvier 2006.

Les détails de l'implémentation "temps réel", des transformations sonores et de la structure de la pièce peuvent être retrouvés dans l'article (Bevilacqua, Rasamimanana, Fléty, Lemouton and Baschet, 2006), présenté à la conférence *NIME* en 2006, disponible en annexe de la thèse.

Florence Baschet travaille actuellement sur une commande de l'IRCAM pour un quatuor *augmenté* où chaque instrumentiste d'un quatuor à corde, en l'occurrence le quatuor Danel, dispose d'un système de captation des mouvements de son archet. Cette formation permet ainsi d'enrichir le paradigme d'interactivité de *BogenLied* en jouant en plus sur les rapports entre les instrumentistes, cherchant parfois l'homogénéité parfois l'opposition dans leurs parties musicales.

Applications pédagogiques

Les systèmes de captations permettent de mesurer différentes variables du mouvement des instrumentistes à corde, en situation de jeu et peuvent ainsi leur donner une image de leur performance. Il peut être intéressant de s'interroger sur l'éventuelle utilisation de tels systèmes dans un contexte pédagogique, où la métaphore du miroir est couramment utilisée : par exemple, le miroir du professeur imitant l'élève, ou l'enregistrement audio, miroir sonore pour l'élève, ou tout simplement le véritable miroir dans lequel l'élève peut se regarder. Des travaux précédents ont proposé d'intégrer dans le processus d'apprentissage d'un instrument de musique des systèmes de cap-

ture du mouvement. Par exemple, l'activité musculaire dans la main gauche d'un violoniste, mesurée à l'aide d'un électro-myogramme (EMG), peut être utilisée pour détecter les problèmes de tension involontaire, potentiellement nuisibles, comme dans l'article (LeVine and Irvine, 1984). Des capteurs embarqués sur un violon, système directement inspiré des travaux de Askenfelt (Askenfelt, 1989), sont utilisés pour visualiser des paramètres de jeu du violon (Ho, 1991), en les utilisant comme trace de la performance.

Les études menées dans cette thèse ont permis d'affiner les modules de captation du mouvement embarqués sur l'archet des instrumentistes pour les rendre de plus en plus précis et de moins en moins intrusifs, et de ce fait utilisable dans un contexte pédagogique. Elles ont également permis de constituer une expertise à partir de laquelle différents scénarios d'utilisation de cette technologie dans un contexte de pédagogie musicale ont été élaborés. L'un des objectifs du miroir, tel que présenté plus haut, est notamment de susciter une autocritique de la part de l'élève, sur ses propres qualités et défauts. Les travaux réalisés à l'IRCAM déclinent cette image du miroir sous différentes formes. La sonification des données gestuelles consiste à traduire de manière sonore les données mesurées par différentes associations (*mappings*) avec des processus de synthèse sonore. Cette approche peut être utilisée pour focaliser l'attention de l'élève sur un problème précis tout en gardant un contexte musical. Par exemple, l'une de mes élèves de violon, travaillant une transcription pour violon du Prélude de la 1ère suite de Bach pour violoncelle, avait du mal à trouver un équilibre dans le bariolage⁷ des différentes cordes. L'association d'un son synthétique pour chaque changement de cordes, détectés à l'aide d'un seuil statique sur la valeur des gyroscopes, permet alors de rendre sonore un geste qui ne l'est pas traditionnellement et de produire dans le cas de mon élève, un motif musical plus ou moins régulier. Pendant une phase de travail, mon élève peut écouter ce motif et modifier la coordination de ses mouvements pour donner par exemple plus de poids aux cordes graves.

Une autre réalisation faite à l'IRCAM permet de créer un "double" de l'élève. Dans le contexte de musique de chambre, par exemple un duo pour

⁷changements rapides sur plusieurs cordes

deux violons, il peut être intéressant pour un élève travaillant seul d'entendre toutes les voix pendant qu'il joue. Le suivi de geste (Bevilacqua et al., 2007) permet de résoudre ce problème tout en laissant une liberté d'interprétation à l'élève : en suivant les mouvements de l'élève jouant une première voix, le système peut jouer de manière synchrone la deuxième voix. Toutefois, dans ce schéma, l'élève contrôle complètement le déroulement de la partie jouée par l'ordinateur, situation différente de celle en musique de chambre où les interprètes généralement écoutent et dirigent à tour de rôle. Néanmoins, ce "double" permet à l'élève de jouer avec la(les) autre(s) voix de la pièce, lui permettant ainsi d'en avoir une connaissance plus approfondie. La technologie est alors utilisée pour développer l'écoute de l'élève, aspect essentiel en pédagogie musicale.

Ces applications à la pédagogie musicale sont notamment utilisées dans le cadre du projet européen I-MAESTRO.

Chapitre 4

Directions futures et conclusion

Cette section présente quelques directions futures de recherche, pendantes aux travaux présentés dans ce manuscrit, notamment applicables à moyen terme. Des éventuelles applications des modèles de profils de vitesse aux recherches sur le suivi de performance et sur le contrôle de processus de synthèse sonore sont proposées. De plus, des directions de recherche dans la continuité des travaux rapportés dans cette thèse sont présentées, notamment concernant l'étude des interactions entre interprètes pour la définition de systèmes d'interaction entre une machine et un(des) interprète(s).

4.1 Suivi de performance

Les travaux menés autour du *suivi de partition* et du *suivi de geste* (Bevilacqua and Muller, 2005; Bevilacqua et al., 2007) à l'IRCAM permettent de suivre en temps-réel l'exécution d'une performance apprise par une machine. Dans un contexte d'interaction entre la machine et l'interprète, il peut être intéressant que la machine non seulement suive le musicien mais qu'elle apprécie également les différences d'interprétation du musicien. Les modèles présentés dans l'article IV peuvent apporter une aide en donnant des informations sur la qualité des mouvements effectués.

Par exemple, à l'aide de la partition musicale, il est possible d'imaginer reconstruire une séquence de mouvements d'archet en utilisant les modèles :

les durées T sont déterminées par la valeurs des notes, les accélérations et longueurs d’archet calculés en fonction des nuances et des phrasés. Une telle *partition de geste*, construite à partir de la partition musicale, peut ensuite être suivie. En captant les mouvements de l’archet des musiciens en temps-réel, il est alors possible d’apprécier, en temps-réel, les différences d’interprétations en utilisant par exemple la corrélation entre les données prédites et les données enregistrées. Cette approche permet alors la machine de devenir ”sensible” aux gestes des instrumentistes à cordes.

4.2 Contrôle de la synthèse sonore

Malgré la richesse de leurs rendus, les algorithmes de synthèse sonore par modèles physiques souffrent de leur complexité de contrôle : produire un son peut s’avérer être aussi compliqué que sur un véritable instrument. Dans ce contexte, il peut alors être intéressant de définir des modèles de ”gestes” de contrôle permettant non seulement de produire un son mais également de créer différentes expressivités comme réalisé dans (Demoucron and Caussé, 2007). L’article IV propose des modèles de profils de vitesse de coups d’archet satisfaisant différentes contraintes physiques et musicales. Ces modèles permettent notamment de générer différentes classes de profils de vitesse avec peu de paramètres. De plus, les paramètres sont faciles d’interprétation (temps, longueur, accélération) et leur influence sur les profils de vitesse est relativement intuitive. Un utilisateur peut alors synthétiser simplement différents profils ”réalistes” de vitesse. Par exemple il peut sélectionner une classe de profil de vitesse (e.g. le modèle trapézoïdal *Trap*) et produire un son plus ou moins long en changeant directement la valeur de T .

D’autre part, la possibilité de synthétiser une séquence de gestes à partir d’une partition musicale, présentée dans le paragraphe précédent, peut être utilisée dans ce cas pour faire ”jouer” un modèle physique de synthèse sonore : l’ordinateur dispose effectivement d’une partition, d’un instrument de musique (modèle physique) et d’une manière de le contrôler pour exécuter la partition (modèles de gestes).

De nombreuses pistes restent encore à explorer. Il est par exemple possible d’imaginer que les modèles soient définis avec des niveaux de contrôle plus ou moins avancés : un premier niveau générant des formes simples comme celles présentées dans ce manuscrit, et un niveau avancé où les formes simples peuvent être modifiées par des effets musicaux tels que des accents, des *crescendo* ou des phrasés. Des analyses plus poussées des mesures de pression de l’archet permettront également de définir des modèles de pression correspondant aux profils de vitesses synthétisés.

4.3 Interactions entre interprètes

L’étude du contrôle de l’archet des instrumentistes en situation de jeu représente une étape fondamentale pour la définition de systèmes interactifs impliquant un instrumentiste (e.g. un violoniste) et une machine. Dans ce contexte, une prochaine étape se situe sans doute dans l’étude des mouvements des musiciens en situation d’interaction. Une situation particulièrement riche peut être celle de la musique d’ensemble.

Par exemple, dans un quatuor à cordes, certaines situations demandent une unité sonore de la part de l’ensemble et d’autres mettent les membres du quatuor en opposition. Dans le premier cas, les instrumentistes cherchent alors à produire un son homogène, propre à l’ensemble. Cet objectif se traduit par un travail spécifique sur les *gestes* musicaux à utiliser en rapport avec l’exécution de la pièce (Quatuor Danel, discussion privée) : cela concerne notamment les longueurs d’archet, les places d’archet, la manière d’effectuer les modes de jeu. Ce travail consiste alors à concilier l’individualité de chacun, jouant sa partie, dans la définition d’un ensemble donnant le son du quatuor. Dans le deuxième cas, certains membres du quatuor doivent se détacher de l’ensemble et mener des dialogues musicaux. Les *gestes* peuvent dans ce cas être amplifiés de manière à rendre ses intentions particulièrement explicites aux autres membres. L’étude des réalisations de ce double objectif par les musiciens peut apporter des éléments de réflexion intéressants dans le contexte des systèmes d’interactions musicales.

4.4 Conclusion

Cette thèse explore les possibilités définies par les contraintes acoustiques et biomécaniques, inhérentes aux mouvements dédiés à la production sonore sur les instruments à cordes frottées. Ce manuscrit présente les résultats suivants :

- . L'étude de ces mouvements dans des contextes musicaux, impliquant différents modes de jeux et différents tempos, permet de caractériser certains aspects des stratégies utilisées par les instrumentistes à des fins expressives.
- . Plusieurs types de segmentation ont été proposées : l'étude à la fois des coups d'archet (tirer - pousser) et des transitions entre coups d'archet offre des vues complémentaires sur le contrôle et la maîtrise de l'archet des instrumentistes.
- . Les modèles définis à partir de minimisations de coût de performance permettent également de regrouper plusieurs coups d'archet en une même unité.
- . L'étude du mouvement d'archet des instrumentistes à l'aide de modèles prenant en compte différentes optimisations de coût de performance permet de mettre en évidence des comportements d'anticipation et de coarticulation dans les mouvements de l'archet.

Ces études peuvent être appliquées à la pédagogie musicale et à la création d'oeuvre de musique contemporaine, créant ainsi des synergies art-recherche enrichissantes pour ces deux domaines.

Bibliographie

- Arfib, D., Couturier, J., Kessous, L., and Verfaillie, V. (2002). Strategies of mapping between gesture data and synthesis model parameters using perceptual spaces. *Organized Sound*, 7(2) :127–144.
- Askenfelt, A. (1986). Measurement of bow motion and bow force in violin playing. *The Journal of the Acoustical Society of America*, 80(4).
- Askenfelt, A. (1989). Measurement of the bowing parameters in violin playing. ii : bow-bridge distance, dynamic range, and limits of bow force. *The Journal of the Acoustical Society of America*, 86(2).
- Askenfelt, A. (1993). Observations on the transient components of the piano tone. In *Proceedings of the Stockholm Music Acoustics Conference (SMAC)*, volume 79, pages 297–301.
- Assayag, G., Bloch, G., Chemillier, M., Cont, A., and Dubnov, S. (2006). Omax brothers : A dynamic topology of agents for improvisation learning. In *ACM Multimedia Workshop on Audio and Music Computing for Multimedia*.
- Assayag, G., Dubnov, S., and Delerue, O. (1999). Guessing the composer’s mind : applying universal prediction to musical style. In *Proceedings of the International Computer Music Conference*.
- Baader, A. P., Kazennikov, O., and Wiesendanger, M. (2005). Coordination of bowing and fingering in violin playing. *Cognitive brain research*, 23(2-3) :436–443.

- Badeau, R. (2005). *Méthodes à haute résolution pour l'estimation et le suivi de sinusoïdes modulées. Application aux signaux de musique*. PhD thesis, Ecole Nationale Supérieure des Télécommunications.
- Badeau, R., Richard, G., and David, B. (2005). Fast adaptive esprit algorithm. In *IEEE Workshop on Statistical Signal Processing SSP'05*, Bordeaux, France.
- Balasubramaniam, R., Wing, A. M., and Daffertshofer, A. (2004). Keeping with the beat : movement trajectories contribute to movement timing. *Experimental Brain Research*, 159(1) :129–134.
- Barthet, M., Kronland-Martinet, R., and Ystad, S. (2006). Consistency of timbre patterns in expressive music performance. In *Proceedings of the Digital Audio Effects Conference (DAFx)*.
- Bell-Berti, F. and Harris, K. (1981). Temporal patterns of coarticulation : lip rounding. *The Journal of the Acoustical Society of America*.
- Bevilacqua, F., Guedy, F., Fléty, E., Leroy, N., and Schnell, N. (2007). Wireless sensor interface and gesture-follower for music pedagogy. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
- Bevilacqua, F. and Muller, R. (2005). A gesture follower for performing arts. In *Proceedings of the International Gesture Workshop*.
- Bevilacqua, F., Rasamimanana, N., and Schnell, N. (2006a). *L'inouï n°2, Revue de l'IRCAM*, chapter Interfaces gestuelles, captation du mouvement et création artistique. Léo Scheer.
- Bevilacqua, F., Rasamimanana, N. H., Fléty, E., Lemouton, S., and Baschet, F. (2006b). The augmented violin project : research, composition and performance report. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.

- Bobick, A. F. and Wilson, A. D. (1997). A state-based approach to the representation and recognition of gesture. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(12) :1325–1337.
- Bril, B. (2002). Apprentissage et contexte. *Intellectica*, 2(35) :251 – 268.
- Cadoz, C. (1988). Instrumental gesture and musical composition. In *Proceedings of the International Computer Music Conference*.
- Cadoz, C. and Wanderley, M. (2000). *Trends in Gestural Control of Music*, chapter Gesture - Music. Ircam - Centre Pompidou.
- Camurri, A., De Poli, G., Friberg, A., Leman, M., and Volpe, G. (2005). The mega project : analysis and synthesis of multisensory expressive gesture in performing art applications. *The Journal of New Music Research*, 34(1) :5–21.
- Chadabe, J. (1997). *C. Chafe's Celletto description in Electric Sound : The Past and Promise of Electronic Music*, chapter 8, pages 226–227. Prentice Hall.
- Chafe, C. (1989). Simulating performance on a bowed instrument. In *Current directions in computer music research*, pages 185 – 198. MIT Press.
- Chafe, C. (1990). Pulsed noise in self-sustained oscillations of musical instruments. In *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, volume 2, pages 1157–1160.
- Cont, A. (2008). Anticipatory score following. (*in preparation*).
- Cremer, L. (1984). *The Physics of the Violin*. The MIT Press.
- Dahl, S. (2000). The playing of an accent - preliminary observations from temporal and kinematic analysis of percussionists. *Journal of New Music Research*, 29 :225–233.
- Dahl, S. (2005). *On the beat : Human movement and timing in the production and perception of music*. PhD thesis, KTH.

- Dahl, S., Bevilacqua, F., Bresin, R., Clayton, M., Hansen, K., Leante, L., Poggi, I., Rasamimanana, N., and Wanderley, M. (2008). *Gesture in performance, chapter in the ConGAS book*. (in preparation).
- Dahl, S. and Friberg, A. (2004). *Expressiveness of Musician's Body Movements in Performances on Marimba, Lecture Notes in Computer Science*, volume 2915, pages 479–486. Springer Verlag.
- Dahl, S. and Friberg, A. (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception*, 24(5) :433–454.
- Dalla Bella, S. and Palmer, C. (2004). Tempo and dynamics in piano performance : the role of movement amplitude. In Lipscomb, S. D., Ashley, R., Gjerdingen, R., and Webster, P., editors, *Proceedings of the International Conference on Music Perception and Cognition*, pages 256–257.
- Dannenbergh, R. (1984). Tracking a live performer. In *Proceedings of the International Computer Music Conference*.
- David, B., Emiya, V., Badeau, R., and Grenier, Y. (2006). Harmonic plus noise decomposition : time-frequency reassignment versus a subspace based method. In *120th AES Convention*, Paris, France.
- De Poli, G., Rodà, A., and Vidolin, A. (1998). Note-by-note analysis of the influence of expressive intentions and musical structure in violin performance. *The Journal of New Music Research*, 27(3) :293–321.
- Demoucron, M. (2007). Measurement of bow pressure in real performance. (*in preparation*).
- Demoucron, M., Askenfelt, A., and Caussé, R. (2006). Mesure de la "pression d'archet" des instruments à cordes frottées : Application à la synthèse sonore. In *Actes du VIIIème Congrès Français d'Acoustique, Tours*.
- Demoucron, M. and Caussé, R. (2007). Sound synthesis of bowed string instruments using a gesture based control of a physical model. In *Proceedings of the International Symposium on Musical Acoustics*.

- Dobrian, C. and Bevilacqua, F. (2003). Gestural control of music using the vicon 8 motion capture system. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
- Engel, K., Flanders, M., and Soechting, J. (1997). Anticipatory and sequential motor control in piano playing. *Experimental Brain Research*, 113 :189–199.
- Flammer, A. and Tordjman, G. (1988). *Le Violon*. J.-C. Lattès and Salabert.
- Fléty, E., Leroy, N., Ravarini, J.-C., and Bevilacqua, F. (2004). Versatile sensor acquisition system utilizing network technology. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
- Freed, A., Wessel, D., Zbyszynski, M., and Uitti, F. M. (2006). Augmenting the cello. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
- Friberg, A., Sundberg, J., and Frydén, L. (2000). Music from motion : Sound level envelopes of tones expressing human locomotion. *The Journal of New Music Research*, 29(3) :199–210.
- Fuchs, A., Jirsa, V. K., Haken, H., and Kelso, J. A. S. (1996). Extending the hkb model of coordinated movement to oscillators with different eigenfrequencies. *Biological cybernetics*, 74(1) :21–30.
- Galamian, I. (1999). *Enseignement et technique du violon*. Van de Velde.
- Gibet, S., Courty, N., and Kamp, J.-F., editors (2006). *Gesture in Human-Computer Interaction and Simulation, 6th International Gesture Workshop, GW 2005, Berder Island, France, May 18-20, 2005, Revised Selected Papers*, volume 3881 of *Lecture Notes in Computer Science*. Springer.
- Goasdoué, R. (2002). Recherches sur l'apprentissage musical instrumental. *Journal de Recherche en Education Musicale*, 1(1) :34–54.

- Godøy, R. I. (2004). Gestural imagery in the service of musical imagery. In Camurri, A. and Volpe, G., editors, *Lecture Notes in Artificial Intelligence, LNAI 2915*, pages 55–62. Springer Verlag.
- Goebl, W., Bresin, R., and Galembo, A. (2004). Once again : The perception of piano touch and tone. can touch audibly change piano sound independently of intensity? In *Proceedings of the International Symposium on Musical Acoustics*.
- Goebl, W., Bresin, R., and Galembo, A. (2005). Touch and temporal behavior of grand piano actions. *The Journal of the Acoustical Society of America*, 118(2) :1154–1165.
- Goodwin, M. (1996). Residual modeling in music analysis-synthesis. In *Proc. IEEE ICASSP*, pages 1005–1008, Atlanta, GA.
- Goto, S. (1999). The aesthetics and technological aspects of virtual musical instruments : The case of the superpolm midi violin. *Leonardo Music Journal*, 9 :115–120.
- Goto, S. and Suzuki, T. (2004). The case study of application of advanced gesture interface and mapping interface, - virtual musical instrument le superpolm and gesture controller bodysuit. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
- Goudeseune, C. (2001). *Composing with parameters for synthetic instruments*. PhD thesis, University of Illinois Urbana-Champaign.
- Goudeseune, C., Garnett, G., and Johnson, T. (2001). Resonant processing of instrumental sound controlled by spatial position. In *Proceedings of the Workshop on New Interfaces for Musical Expression (NIME)*.
- Grey, J. M. (1975). *An exploration of musical timbre using computer-based techniques for analysis, synthesis and perceptual scaling*. PhD thesis, Stanford University.

- Guettler, K. (2002a). *The Bowed String : On the Development of Helmholtz Motion and On the Creation of Anomalous Low Frequencies*. PhD thesis, KTH.
- Guettler, K. (2002b). On the creation of the helmholtz motion in bowed strings. *Acta Acustica*, 88.
- Guettler, K. (2004a). Applications of the bluestein filter in bowed-string. In *Proceedings of Frontiers of Research on Speech and Music*.
- Guettler, K. (2004b). Looking at starting transients and tone coloring of the bowed string. In *Proceedings of Frontiers of Research on Speech and Music*.
- Guettler, K. and Askenfelt, A. (1997). Acceptance limits for the duration of pre-helmholtz transients in bowed string attacks. *The Journal of the Acoustical Society of America*, 101(5) :2903–2913.
- Guettler, K., Schoonderwaldt, E., and Askenfelt, A. (2003). Bow speed or bowing position - which one influences spectrum the most? In *Proceedings of the Stockholm Music Acoustics Conference (SMAC)*.
- Haken, H., Kelso, J. A. S., and Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological cybernetics*, 51(5) :347–356.
- Henrich, N. (2006). Mirroring the voice from garcia to the present day : Some insights into singing voice registers. *Logopedics Phoniatrics Vocology*, 31 :3–14(12).
- Hikichi, T., Osaka, N., and Itakura, F. (2004). Sho-so-in : Control of a physical model of the sho by means of automatic feature extraction from real sounds. *The Journal of New Music Research*, 33(4) :355–365.
- Ho, T. (1991). A computer-assisted approach to the teaching of violin tone production. *SIGCUE Outlook*, 21(2) :73–83.

- Hogan, N. (1984). An organizing principle for a class of voluntary movements. *The Journal of Neuroscience*, 4(11) :2745–2754.
- Hogan, N. and Sternad, D. (2007). On rhythmic and discrete movements : reflections, definitions and implications for motor control. *Experimental Brain Research*, 181 :13–30.
- Hunt, A. and Wanderley, M. (2002). Mapping performer parameters to synthesis engines. *Organized Sound*, 7(2).
- Hunt, A., Wanderley, M., and Paradis, M. (2003). The importance of parameter mapping in electronic instrument design. *The Journal of New Music Research*, 32(4).
- Jensenius, A., Kvifte, T., and Godøy, R. I. (2006). Towards a gesture description interchange format. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
- Jordà, S., Kaltenbrunner, M., Geiger, G., and Bencina, R. (2005). The reactable. In *Proceedings of the International Computer Music Conference*.
- Kapur, A., Davidson, P., Cook, P., Schloss, A., and Driessen, P. (2005). Preservation and extension of traditional techniques : Digitizing north indian performance. *The Journal of New Music Research*.
- Kapur, A., Lazier, A., Davidson, P., Wilson, R., and Cook, P. (2004a). The electronic sitar controller. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
- Kapur, A., Tzanetakis, G., and Driessen, P. (2004b). Audio-based gesture extraction on the esitar controller. In *Proceedings of the Digital Audio Effects Conference (DAFx)*.
- Kereliuk, C., Scherrer, B., Verfaillie, V., Depalle, P., and Wanderley, M. (2007). Indirect acquisition of fingerings on harmonic notes on the flute. In *Proceedings of the International Computer Music Conference*.

- Koehly, R., Wanderley, M., and Curtil, D. (2006). Paper fsrs and latex/fabric traction sensors : methods for the development of home-made touch sensors. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
- Laroche, J. (1989). A new analysis/synthesis system of musical signals using prony's method-application to heavily damped percussive sounds. In *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, volume 3, pages 2053–2056.
- Leman, M. (2007). *Embodied music cognition and mediation technology*. The MIT Press.
- Leroy, N., Fléty, E., and Bevilacqua, F. (2006). Reflective optical pickup for violin. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
- Letz, S. (1998). *recherches et applications en informatiques musicales*, chapter 13 : les normes MIDI et MIDIFiles. M. Chemilier and F. Pachet.
- LeVine, W. and Irvine, J. (1984). In vivo emg biofeedback in violin and viola pedagogy. *Biofeedback and self-regulation*, 9(2).
- Loehr, J. D. and Palmer, C. (2007). Cognitive and biomechanical influences in pianists' finger tapping. *Experimental Brain Research*, 178 :518–528.
- Loog, M. and Duin, R. P. W. (2004). Linear dimensionality reduction via a heteroscedastic extension of lda : The chernoff criterion. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(6).
- Machover, T. and Chung, J. (1989). Hyperinstruments : Musically intelligent and interactive performance and creativity systems. In *Proceedings of the International Computer Music Conference*.
- Maestre, E., Bonada, J., Blaauw, M., Perez, A., and Guaus, E. (2007). Acquisition of violin instrumental gestures using a commercial emf tracking device. In *Proceedings of the International Computer Music Conference*.

- Malloch, J., Birnbaum, D., Sinyor, E., and Wanderley, M. (2006). Towards a new conceptual framework for digital musical instruments. In *Proceedings of the Digital Audio Effects Conference (DAFx)*.
- Mancini, M., Bresin, R., and Pelachaud, C. (2007). A virtual head driven by music expressivity. *IEEE Transactions on Audio, Speech and Language Processing*, 15(6).
- Masri, P. and Bateman, A. (1996). Improved modelling of attack transients in music analysis-resynthesis. In *Proc. of the International Computer Music Conference*.
- McAdams, S., Winsberg, S., de Soete, G., and Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres : common dimensions, specificities and latent subject classes. *Psychological Research*, 58 :177–192.
- McIntyre, M. E., Schumacher, R. T., and Woodhouse, J. (1981). Aperiodicity in bowed string motion. *Acustica*, 49(1) :13–32.
- McIntyre, M. E., Schumacher, R. T., and Woodhouse, J. (1984). A parametric study of the bowed string : the violinist’s menagerie. *Journal of Catgut Acoustical Society*, 42 :18 – 21.
- McMillen, K. (1994). Zipi - origins and motivations. *The Computer Music Journal*, 18(4) :47–51.
- Menuhin, Y. (1973). *L’Art de jouer du violon*. Buchet / Chastel.
- Moore, R. (1988). The dysfunctions of midi. *The Computer Music Journal*, 12(1).
- Morin, E. (1999). Captation de modes de jeu instrumentaux : le cas du violoncelle. Technical report, Ircam.
- Morin, E. (2000). Analyse des coups d’archet du violoncelliste avec le dispositif digibow. Technical report, Ircam.

- Nelson, W. L. (1983). Physical principles for economies of skilled movements. *Journal Biological Cybernetics*, 46(2) :135–147.
- Nichols, C. (2002). The vbow : Development of a virtual violin bow haptic human-computer interface. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
- Orio, N. (1999). The timber space of the classical guitar and its relationship with the plucking techniques. In *Proceedings of the International Computer Music Conference*.
- Orio, N. and Dechelle, F. (2001). Score following using spectral analysis and hidden markov models. In *Proceedings of the International Computer Music Conference*.
- Orio, N., Schnell, N., and Wanderley, M. (2001). Evaluation of input devices for musical expression : Borrowing tools from hci. *The Computer Music Journal*, 26(3) :62–76.
- Ostry, D., Cooke, J., and Munhall, K. (1987). Velocity curves of human arm and speech movements. *Experimental Brain Research*, 68 :37–46.
- Overholt, D. (2005a). The overtone violin. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
- Overholt, D. (2005b). The overtone violin : a new computer music instrument. In *Proceedings of the International Computer Music Conference*.
- Pachet, F. (2002). Playing with virtual musicians : the continuator in practice. *IEEE Multimedia*.
- Palmer, C. (2001). Movement amplitude and tempo change in piano performance. *The Journal of the Acoustical Society of America*, 115 :2590–2590.
- Paradiso, J. A. and Gershenfeld, N. (1997). Musical applications of electric field sensing. *Computer Music Journal*, 21(2).

- Peeters, G. (2004). A large set of audio features for sound description (similarity and classification). Technical report, Ircam.
- Peiper, C., Warden, D., and Garnett, G. (2003). An interface for real-time classification of articulations produced by violin bowing. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
- Penttinen, H. (2006). On the dynamics of the harpsichord and its synthesis. In *Proceedings of the Digital Audio Effects Conference (DAFx)*.
- Perkell, J., Zandipour, M., Matthies, M., and Lane, H. (2002). Economy of effort in different speaking conditions. i. a preliminary study of intersubject differences and modeling issues. *The Journal of the Acoustical Society of America*, 112(4) :1627–1641.
- Poepel, C. (2004). Synthesized strings for string players. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
- Poepel, C. and Marx, G. (2007). >hot strings sig<. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
- Rabiner, L. R. (1990). A tutorial on hidden markov models and selected applications in speech recognition. In Waibel, A. and Lee, K.-F., editors, *Readings in Speech Recognition*, pages 267–296. Kaufmann, San Mateo, CA.
- Rank, E. (1999). A player model for midi control of synthetic bowed strings. In *Diderot Forum on Mathematics and Music*.
- Raphael, C. (1999). Automatic segmentation of acoustic musical signals using hidden markov models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(4) :360–370.
- Rasamimanana, N. H. (2004). Gesture analysis of bow strokes using an augmented violin. Master’s thesis, Université Paris 6, IRCAM CNRS STMS.

- Rasamimanana, N. H., Bernardin, D., Wanderley, M., and Bevilacqua, F. (2007a). String bowing gestures at varying bow stroke frequencies : A case study. In *Advances in Gesture-Based Human-Computer Interaction and Simulation*, volume 5085 of *Lecture Notes in Computer Science*, pages 216–226. Springer Verlag.
- Rasamimanana, N. H. and Bevilacqua, F. (2008). Effort-based analysis of bowing movements : evidence of anticipation effects. *The Journal of New Music Research (accepted for publication)*.
- Rasamimanana, N. H., Fléty, E., and Bevilacqua, F. (2006). Gesture analysis of violin bow strokes. In *Gesture in Human-Computer Interaction and Simulation*, volume 3881 of *Lecture Notes in Computer Science / Artificial Intelligence (LNAI)*, pages 145–155. Springer Verlag.
- Rasamimanana, N. H., Kaiser, F., and Bevilacqua, F. (2007b). Transients control of violin players : relationships between bow acceleration and string irregular vibrations. *The Journal of New Music Research (submitted)*.
- Reng, L., Moeslund, T. B., and Granum, E. (2006). Finding motion primitives in human body gestures. In Gibet, S., Courty, N., and Kamp, J.-F., editors, *Lecture Notes in Artificial Intelligence, LNAI 3881*, pages 133–144. Springer Verlag.
- Risset, J.-C. and Van Duyne, S. (1996). Real-time performance interaction with a computer-controlled acoustic piano. *The Computer Music Journal*, 20(1) :62–75.
- Rowe, R. (1992). Machine listening and composing with cypher. *Computer Music Journal*, 16(1) :43 – 63.
- Schelleng, J. (1973). The bowed string and the player. *The Journal of the Acoustical Society of America*, 53(1).
- Schnell, N., Bevilacqua, F., Lyons, M. J., and Tanaka, A., editors (2006). *New Interfaces for Musical Expression, NIME-06, Proceedings, Paris, June*

- 4-8, 2006. IRCAM - Centre Pompidou in collaboration with Sorbonne University.
- Schoner, B., Cooper, C., Douglas, C., and Gershenfeld, N. (1998). Data-driven modelling and synthesis of acoustical instruments. In *Proceedings of the International Computer Music Conference*.
- Schoner, B., Cooper, C., Douglas, C., and Gershenfeld, N. (2000). Cluster-weighted sampling for synthesis and cross-synthesis of violin family instrument. In *Proceedings of the International Computer Music Conference*.
- Schoonderwaldt, E., Guettler, K., and Askenfelt, A. (2003). Effect of the width of the bow hair on the violin string spectrum. In *Proceedings of the Stockholm Music Acoustics Conference (SMAC)*.
- Schoonderwaldt, E., Guettler, K., and Askenfelt, A. (2007). An empirical investigation of bow-force limits in the schelleng diagram. (*submitted*).
- Schoonderwaldt, E., Rasamimanana, N., and Bevilacqua, F. (2006). Combining accelerometer and video camera : Reconstruction of bow velocity profiles. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, pages 200–203.
- Schoonderwaldt, E. and Wanderley, M. (2007). Visualization of bowing gestures for feedback : the hodgson plot. In *Proceedings of the I-Maestro Workshop*.
- Schwarz, D. (2004). *Data-Driven Concatenative Sound Synthesis*. PhD thesis, Université Paris 6.
- Schwarz, D., Orio, N., and Schnell, N. (2004). Robust polyphonic midi score following with hidden markov models. In *Proceedings of the International Computer Music Conference*.
- Serafin, S. (2004). *The sound of friction : real-time models, playability and musical applications*. PhD thesis, Stanford University, CCRMA.

- Serafin, S., Dudas, R., Wanderley, M., and Rodet, W. (1999). Gestural control of a real-time physical model of a bowed string instrument. In *Proceedings of the International Computer Music Conference*.
- Serafin, S. and Young, D. (2003). Bowed string physical model validation through use of a bow controller and examination of bow strokes. In *Proceedings of the Stockholm Music Acoustics Conference (SMAC)*.
- Serra, X. and Smith III, J. O. (1990). Spectral modeling synthesis : A sound analysis/synthesis system based on a deterministic plus stochastic decomposition. *Computer Music Journal*, 14(4) :12–24.
- Shan, G. B. and Visentin, P. (2003). A quantitative three-dimensional analysis of arm kinematics in violin performance. *Medical Problems of Performing Artists*, 18(1) :3–10.
- Styns, F., van Noorden, L., and Leman, M. (2006). Some basic observations on how people move on music and how they relate music to movement. In *Proceedings of the International Music and Gesture Conference*.
- Tanaka, A. (2000). *Trends in Gestural Control of Music*, chapter Musical Performance Practice on Sensor-based Instruments, page 389 to 406. Ircam.
- Tanaka, A. and Knapp, B. (2002). Multimodal interaction in music using the electromyogram and relative position sensing. In *Proceedings of the International Computer Music Conference*.
- Thibault, F., editor (2003). *New Interfaces for Musical Expression, NIME-03, Proceedings, Montréal, May 22-24, 2003*. Faculty of Music, McGill University.
- Traube, C. and D’Alessandro, N. (2005). Vocal synthesis and graphical representation of the phonetic gestures underlying guitar timbre description. In *Proceedings of the Digital Audio Effects Conference (DAFx)*.

- Traube, C. and Depalle, P. (2004). Phonetic gestures underlying guitar timbre description. In *Proceedings of the International Conference on Music Perception and Cognition*.
- Traube, C., Depalle, P., and Wanderley, M. M. (2003). Indirect acquisition of instrumental gesture based on signal, physical and perceptual information. In Thibault (2003), pages 42–48.
- Trueman, D. and Cook, P. R. (1999). Bossa : The deconstructed violin reconstructed. In *Proceedings of the International Computer Music Conference*.
- Turner-Stokes, L. and Reid, K. (1998). Three-dimensional motion analysis of upper limb movement in the bowing arm of string-playing musicians. *Clinical Biomechanics*.
- Van Nort, D., Wanderley, M., and Depalle, P. (2004). On the choice of mappings based on geometric properties. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
- Vercoe, B. (1984). The synthetic performer in the context of live performance. In *Proceedings of the International Computer Music Conference*.
- Verfaillie, V., Quek, O., and Wanderley, M. (2006a). Sonification of musician’s ancillary gestures. In *Proceedings of the International Conference on Auditory Display*.
- Verfaillie, V., Wanderley, M., and Depalle, P. (2006b). Mapping strategies for gestural and adaptive control of digital audio effects. *The Journal of New Music Research*, 35(1) :71–93.
- Verma, T., Levine, S., and Meng, T. (1997). Transient modeling synthesis : a flexible analysis/synthesis tool for transient signals. In *Proceedings of the International Computer Music Conference*.
- Wagennar, R. and van Emmerik, R. (2000). Resonant frequencies of arms and legs identify different walking patterns. *Journal of Biomechanics*, 33 :853–861.

- Waiswiz, M. (1985). The hands, a set of remote midi-controllers. In *Proc. of the International Computer Music Conference*.
- Wanderley, M. (2001). *Performer-Instrument interaction : applications to gestural control of sound synthesis*. Thèse, Université Paris 6.
- Wanderley, M. and Battier, M., editors (2000). *Trends in Gestural Control of Music*. Ircam.
- Wanderley, M. and Depalle, P. (2004). Gestural control of sound synthesis. In *Proceedings of the IEEE*, volume 92, pages 632–644.
- Wanderley, M., Vines, B., Middleton, N., McKay, C., and Hatch, W. (2005). The musical significance of clarinetists’ ancillary gestures : An exploration of the field. *The Journal of New Music Research*, 34(1) :97–113.
- Wessel, D. and Wright, M. (2002). Problems and prospects for intimate musical control of computers. *The Computer Music Journal*, 26(3) :11–22.
- Winold, H., Thelen, E., and Ulrich, B. D. (1994). Coordination and control in the bow arm movements of highly skilled cellists. *Ecological Psychology*, 6(1) :1–31.
- Wolfe, J. (2002). Speech and music, acoustics and coding, and what music might be ‘for’. In *Proceedings of the 7th International Conference on Music Perception and Cognition, Sydney*.
- Woodhouse, J. and Galluzzo, P. M. (2004). The bowed string as we know it today. *Acustica – Acta Acustica*, 90 :579 – 589.
- Yeh, C. and Röbel, A. (2006). Adaptive noise level estimation. In *Proc. Workshop on Computer Music and Audio Technology, Taipei*.
- Young, D. (2003). Wireless sensor system for measurement of violin bowing parameters. In *Proceedings of the Stockholm Music Acoustics Conference (SMAC)*.

- Young, D. (2007). *A Methodology for Investigation of Bowed String Performance Through Measurement of Violin Bowing Technique*. PhD thesis, M.I.T.
- Young, D. and Deshmane, A. (2007). Bowstroke database : a web-accessible archive of violin bowing data. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
- Young, D., Nunn, P., and Vassiliev, A. (2006). Composing for hyperbow : A collaboration between mit and the royal academy of music. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*.
- Young, D. and Serafin, S. (2007). Investigating the performance of a violin physical model. In *Proceedings of the International Computer Music Conference*.

Deuxième partie

Articles

ARTICLE I

Gesture Analysis of Violin Bow Strokes

Nicolas H. Rasamimanana, Emmanuel Fléty, and Frédéric Bevilacqua

IRCAM

1, Place Igor Stravinsky

75004 PARIS

{Nicolas.Rasamimanana, Emmanuel.Flety, Frederic.Bevilacqua}@ircam.fr

Abstract. We developed an "augmented violin", i.e. an acoustic instrument with added gesture capture capabilities to control electronic processes. We report here gesture analysis we performed on three different bow strokes, *Détaché*, *Martelé* and *Spiccato*, using this augmented violin. Different features based on velocity and acceleration were considered. A linear discriminant analysis has been performed to estimate a minimum number of pertinent features necessary to model these bow stroke classes. We found that the maximum and minimum accelerations of a given stroke were efficient to parameterize the different bow stroke types, as well as differences in dynamics playing. Recognition rates were estimated using a kNN method with various training sets. We finally discuss that bow stroke recognition allows to relate the gesture data to music notation, while a bow stroke continuous parameterization can be related to continuous sound characteristics.

Key words: Music, Gesture Analysis, Bow Strokes, Violin, Augmented Instruments

1 Introduction

There is an increasing interest in using gestural interfaces to control digital audio processes. Despite numerous recent achievements ([12]), important ground work on gesture analysis is still necessary for the improvement of such interfaces. We are currently developing various "augmented instruments", i.e. acoustic instruments with added gesture capture capabilities. Such an approach remains remarkably fruitful for the study of gesture in music. As a matter of fact, the use of acoustic instruments in this context allows to apprehend instrumental gesture in a *a priori* defined framework, linked to both a symbolic level, the music notation, and a signal level, the acoustic instrument sound.

One of our current project concerns an "augmented violin", similar to the one developed by D. Young [13]. On a fundamental level, our goal is to build a model of the player's gestures reflecting his/her expressive intentions related to violin playing techniques. Specifically, our aims are to establish the relationships between the captured data, bowing styles and sound characteristics. This

includes the study, on a gestural level, of the variations that occur between different interpretations of a single player or between players. These studies will lead us to the development of real-time analysis tools, enabling an *interpretation feedback*, which includes gesture recognition, and *gesture following*, i.e. the possibility to track a performance with respect to a predefined reference. We believe that both approaches are key to develop novel types of interaction between instrumentalists and computers.

We report in this paper the study of three violin bow strokes (*Détaché*, *Martelé* and *Spiccato*) and the evaluation of their possible recognition. The article is organized as follows. We first present a review of similar works. In section 3, we present the capture system implemented on the violin. In sections 4 and 5, we show results on the parameterization and recognition of bow stroke types. Finally, we conclude in sections 6 and 7 by a discussion of these results and their implications on future work.

2 Related Works

Our concept of "augmented instruments" is similar to the Hyperinstruments developed by T. Machover and collaborators. The idea is to use a traditional instrument and to extend its capabilities by digital means. For example, the HyperCello [7] created in 1991 was conceived as an acoustic cello with added measurements of wrist movement, bow pressure and position, and left hand fingering. More recently, D. Young extended the HyperCello to the violin with the HyperBow [13], [14].

Several other interfaces have been developed based on string instruments for artistic purposes ([6], [4], [11], and [9]). All of these works generally used the sensor signals to directly control sound treatment parameters, such as filters [11] or physical model synthesis [14]. B. Schoner [10] adopted a probabilistic approach to infer, in real time, cello sounds from the gesture input given by the HyperCello.

Very few works actually report an analysis of the signals, specifying the relationships between the data and the instrumentist's performance. Among them, C. Peiper et al. [8] used decision tree techniques to classify violin bow strokes based on motion tracking. We here pursue the approach of analyzing different types of bow strokes, and in particular we propose to estimate invariance and variability of the measured signals.

3 Hardware Design

Hardware developments were designed with the following constraints: compatibility with an acoustic violin, no significant alteration of the instrument, wireless communication, relatively inexpensive. The prototype we built and used in this study is shown on figure 1. Two types of gesture data are measured, using technology similar to the one described in [13]: bow position and bow accelerations.

First, the sensing system can measure the bow-strings contact position along two directions: between tip and frog, between bridge and finger-board. This position is measured via capacity coupling between a resistive tape fixed along the bow and an antenna behind the bridge.

Second, acceleration is sensed thanks to two Analog Device ADXL202 placed at the bow frog. Note that such sensors are sensitive to both gravity, hence inclination, and movement acceleration (generally referred as static and dynamic accelerations). The two accelerometers are fixed to the bow nut in such a way that acceleration is measured in three dimensions: bowing direction, string direction and vertical direction.

The position data, obtained from the antenna behind the bridge, is digitized in 16 bits with a sensor acquisition system developed at IRCAM, *Ethersense* [3]. The acceleration data are sent wirelessly to a RF receiver also connected to the sensor acquisition system. The acceleration dynamic range has been measured to be of 65 dB. All the data are transmitted to Max/MSP through an ethernet connection using the Open Sound Control protocol, at a data rate of 200 Hz. The surplus weight added by the sensing system is actually 15 grams, mainly located at the frog. Although perceptively heavier, the bow is still playable according to professional violinists. A smaller and slightly lighter prototype is currently under development.

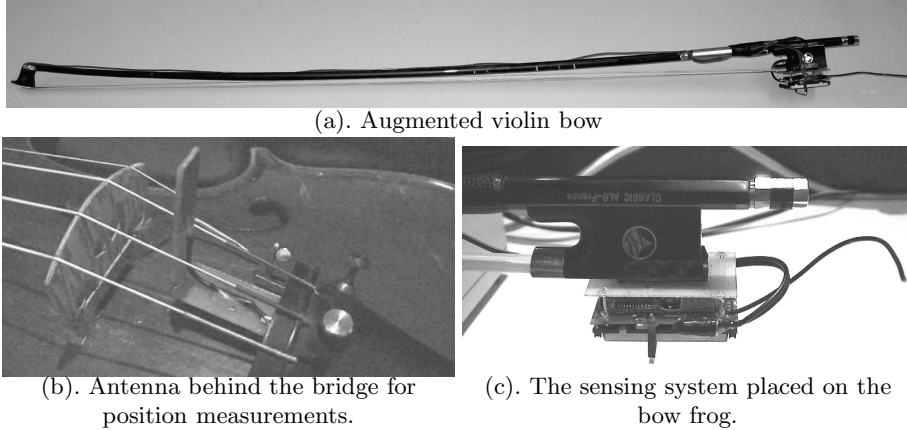


Fig. 1. Pictures of the augmented violin prototype.

4 Gesture Analysis

We studied three standard types of bow strokes (*Détaché*, *Martelé* and *Spiccato*), by focusing the analysis on accelerometer signals in the bowing direction, which contain the essential information.

4.1 Violin Bow Strokes

Here is a brief description of these bow strokes according to [2].

In *Détaché*, the bow linearly goes from tip to frog and inversely from frog to tip. This linear movement must be adapted to the various dynamics. The bow can be used entirely or in fractions.

Martelé requires a violent gesture. The whole arm must be rapid and vigorous: a very sharp, almost percussive attack must be obtained at each extremity of the bow.

Spiccato uses the phalanges suppleness so that the bow can leave the string after each notes. It results in a light and precise sound.

4.2 Data Acquisition

We built a database from recordings of professional and amateur violinists performing scales in the three bow strokes *Détaché*, *Martelé* and *Spiccato*, at two tempi, 60 bpm and 120 bpm, and three dynamics, *pianissimo* (*pp*), *mezzo forte* (*mf*), *fortissimo* (*ff*).

In order to free the accelerometer signals from angle contributions, we asked the violinists to perform scales on one string at a time and recorded scales on every strings. This way, angle contribution is a constant offset and can be subtracted.

We chose in this study to consider individual strokes. We therefore segmented the recorded gesture data using a peak detection algorithm on the acceleration signals. The gesture database is hence constituted of executions of separate notes, played in three different styles, at three dynamics, two tempi, by two different players.

Figure 2 shows an example of data for the three types of bow strokes *mf* and at 60 bpm. We can see that in *Détaché*, bow velocity remains relatively constant after the attack, unlike *Martelé* and *Spiccato*, where the bow must be slowed down. *Martelé* has typically higher absolute acceleration values compared to *Spiccato*. *Martelé* indeed requires more velocity as it is generally performed using a greater length of bow, compared to *Spiccato*, in order to achieve its typical percussive attack.

4.3 Gesture Features

Four parameters are derived from the acceleration and velocity curves to model the bow strokes: a_{max} , a_{min} , v_{max} and v_{min} (first local minimum after v_{max}), as illustrated on figure 3. Bow velocity is computed from the integration of accelerometers signals. These features correspond to a basic parameterization of the velocity curve shape. They can be computed with sufficient precision and without assuming any model for the velocity shape. They allow for the representation of *Détaché*, *Martelé* and *Spiccato* within a four dimensional space.

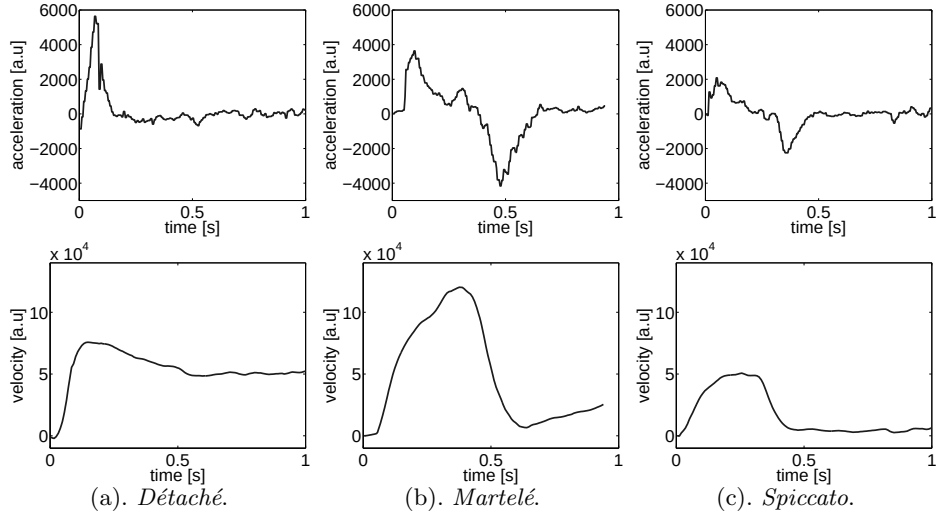


Fig. 2. Acceleration and velocity curves for a single note played in the three styles *Détaché*, *Martelé* and *Spiccato*. Dynamic is *mf* and tempo 60 bpm

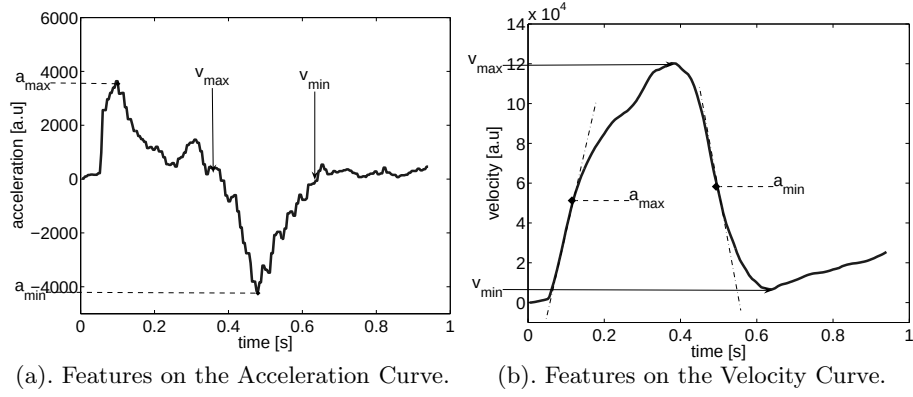


Fig. 3. Illustration of the four features a_{max} , a_{min} , v_{max} and v_{min} (first local minimum after v_{max}) on *Martelé* acceleration and velocity curves.

4.4 Gesture Space

We used Linear Discriminant Analysis (LDA), which maximizes separation between classes, to estimate the dimensionality of the parameterization. LDA on the gesture database, considering three bow strokes classes, indicates that the class scatter matrix only has two significant eigen values. Therefore, the gesture data can be clustered in a bidimensional space, with maximum in-between classes distance.

We actually found that a_{max} and a_{min} , having major contributions in the eigen vectors, are the two most consistent parameters to model bow strokes, as illustrated in figures 4 and 5. As shown on figures 4(a) and 4(b), for a given dynamic, each bow stroke type forms a separate cluster. Moreover, the disposition of these clusters is similar for both players.

Figure 5(a) illustrates the case where different dynamics are considered. The basic clustering structure remains even if overlap occurs. Nevertheless, for each bow stroke types, sub-structure clustering can be observed as detailed in figures 5(b), 5(c) and 5(d). Precisely, each cluster is composed of three sub-clusters, one for each dynamic variations (pp , mf , ff). *Fortissimo* always corresponds to the highest a_{max} and a_{min} values.

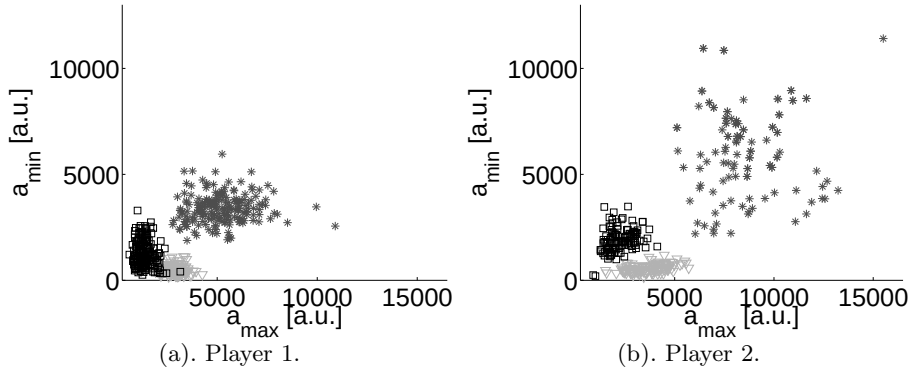
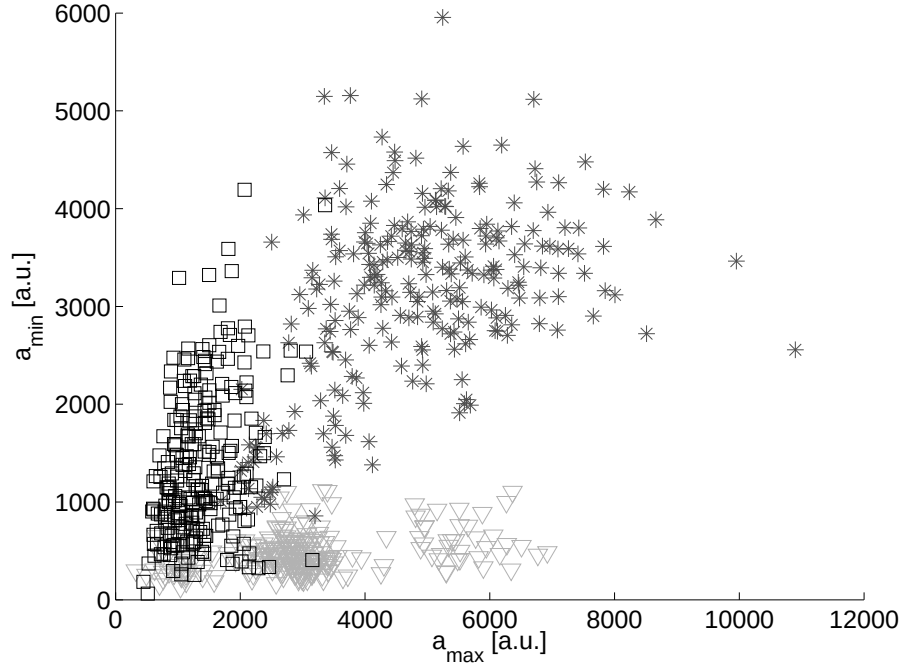


Fig. 4. Bow Strokes Feature Space (Player Detail). Each point corresponds to a single bow stroke. Fig (a) and (b) show the feature space for each player, at a same dynamic (mf) and tempo (60 bpm). Legend is *Detaché* = ∇ , *Martelé* = $*$, and *Spiccato* = $\square$.

5 Gesture Recognition

We further evaluate the ability of recognizing bow stroke using kNN with a_{max} and a_{min} . Three different test scenarios were chosen.

First, we defined three classes, corresponding to the three types of bow strokes. The whole database, i.e. mixing two players, three dynamics and two tempi, is randomly divided into two parts (one-fourth and three-fourths). The



(a). Bow stroke feature space mixing three dynamics for one player and one tempo (60 bpm). Legend is *Détaché* = ∇ , *Martelé* = $*$, and *Spiccato* = $\square$.

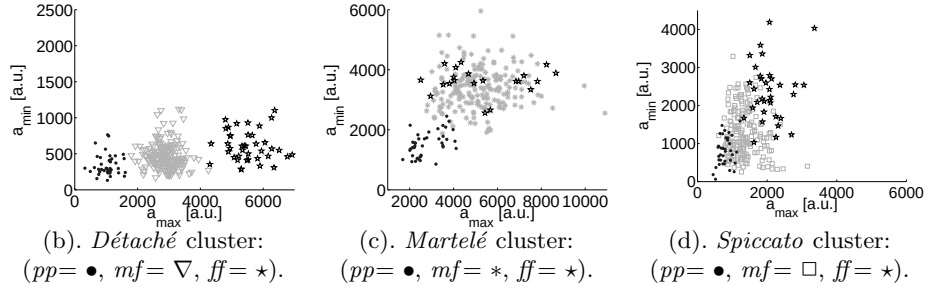


Fig. 5. Bow Strokes Feature Space (Dynamic Detail). Each point corresponds to a single bow stroke. Fig (a) plots all the features points for one player, at one tempo and at three dynamics. The three bow strokes appear in clusters. Fig (b), (c) and (d) show the detail for each bow stroke cluster: three sub-clusters corresponding to the three dynamics can be seen.

quarter of the database, i.e. 320 points, serves as a reference and the remaining three quarter, i.e 1000 points, is used to evaluate the recognition rate. For each test point, vote is done according to the most represented type of bow stroke in the 10 nearest neighbors. Table 1 shows the recognition percentage for this first setup.

In the second scenario, we considered the same three classes but cross-tested the players data: one served as reference for the other. The results are reported in table 2.

In the third scenario, we considered each variation of dynamics as a separate class. Thus, nine reference classes, i.e. three types of bow strokes times three nuances for a single player, are tested. This time, two-thirds of the database are used as a reference where each of the nine classes is represented. Table 3 shows the recognition results. For each line, first column is the class of the tested points and the other columns give the percentages of recognition for the nine classes.

Table 1. kNN recognition results (Test scenario 1). Database is mixing 2 players, 3 nuances and 2 tempi. Three classes considered.

<i>Test\Ref</i>	<i>Détaché</i>	<i>Martelé</i>	<i>Spiccato</i>
<i>Détaché</i>	96.7%	1.3%	2.0%
<i>Martelé</i>	1.0%	85.8%	13.2%
<i>Spiccato</i>	6.0%	5.0%	89.0%

Table 2. kNN recognition results (Test scenario 2). Database is mixing 1 player (*Pl1*), 1 nuance (*mf*) and 1 tempo (60bpm). Test points from other player (*Pl2*), same nuance and tempo. Three classes considered.

Ref		<i>Pl1</i>		
Test \		<i>Det</i>	<i>Mar</i>	<i>Spi</i>
<i>Pl2</i>	<i>Det</i>	100.0%	0.0%	0.0%
	<i>Mar</i>	0.0%	100.0%	0.0%
	<i>Spi</i>	6.3%	25.0%	68.7%

6 Discussion

For recognition, three test scenarios were elaborated. The first two scenarios yields high recognition rates. This shows that the three bow strokes are efficiently characterized by the features (a_{min} , a_{max}), even mixing data of two players, different dynamics and tempi, and with a relatively low number of reference

Table 3. kNN recognition results (Test scenario 3). Database is mixing 1 player (*Pl1*), 3 nuances, 1 tempo (60bpm). Nine classes considered (3 bow strokes x 3 nuances).

Ref		pp			mf			ff		
Test \		Det	Mar	Spi	Det	Mar	Spi	Det	Mar	Spi
pp	Det	100.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %
	Mar	0.0 %	78.6 %	0.0 %	0.0 %	0.0 %	7.1 %	0.0 %	0.0 %	14.3 %
	Spi	23.1 %	0.0 %	76.9 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %
mf	Det	0.0 %	9.5 %	0.0 %	90.5 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %
	Mar	0.0 %	0.0 %	0.0 %	0.0 %	95.8 %	0.0 %	4.2 %	0.0 %	0.0 %
	Spi	0.0 %	35.3 %	0.0 %	0.0 %	0.0 %	64.7 %	0.0 %	0.0 %	0.0 %
ff	Det	0.0 %	0.0 %	0.0 %	6.7 %	0.0 %	0.0 %	93.3 %	0.0 %	0.0 %
	Mar	0.0 %	0.0 %	0.0 %	0.0 %	85.7 %	0.0 %	0.0 %	14.3 %	0.0 %
	Spi	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	49.9 %	0.0 %	0.0 %	50.1 %

data, i.e. one-fourth of the data. Moreover, the cross-player test done in the second scenario confirms the features invariance properties. In this well defined playing situation (scales), our results thus show that the chosen features can be directly related to a music notation level.

In the third scenario, the recognition performances are reduced in some cases. Even with a high proportion of data as reference (two-thirds), confusions occur for example between *Spiccato mf* and *Martelé pp*. However, such confusions are informative as they illustrate actual similarities in bow stroke gestures, when mixing different dynamics. Precisely, from our results, the following different classes, *Spiccato mf*, *Martelé pp* and *Détaché mf*, share similar features, which was actually found to be consistent from the viewpoints of violinists. This shows the limits of recognition approaches since frontiers between classes are not always well defined perceptively.

Furthermore, points that are close in the gesture feature space (figure 5(a)) share similar sound characteristics, e.g. *Martelé pp*, *Détaché mf* and *Spiccato ff*. Consequently, it is perceptually more coherent to characterize bow strokes with a continuous parameterization, using for example a_{max} and a_{min} : such parameters can indeed be related to continuous sound characteristics and/or perceptual features of the listener. It is important to note that a continuous parameterization enables both the recognition of bowing styles and the characterization of hybrid bow strokes.

The results of the study also show that bow acceleration is a parameter of major influence to characterize the different ways of bowing. This comes in complement to acoustic studies on the violin, notably by A. Askenfelt [1] and K. Guettler [5], having already demonstrated the influence of bow acceleration values on the establishments of a Helmholtz regime. It will be interesting to relate the different bowing styles to the number of nominal periods elapsing before Helmholtz triggering occurs, as described in [5].

7 Conclusion and Perspectives

Our goal was to study three different bow strokes, *Détaché*, *Martelé* and *Spiccato*, based on gesture data. After considering basic features based on velocity and acceleration curves, we found that a_{max} and a_{min} provided a pertinent parameterization of these bow strokes. In particular, these parameters enable the recognition of bow stroke types (even in the case of two different players). When considering a higher number of classes including dynamics, we noted typical confusions, consistent with perceptual point of views of violin players and listeners. In summary, our gesture analysis was based on two complementary approaches: recognition and gesture parameterization. Recognition allows us to relate gesture data to music notation, while continuous parameterization of bow strokes could be related to continuous sound characteristics. The detailed relationship between gesture data and sound parameters will be the object of a future study. Moreover, we will investigate other type of parameterizations of the velocity and acceleration that should account for finer characterization of bow strokes. Other parameters such as bow force on strings, pointed by acoustic studies as an influential parameter on sound, will also be considered.

8 Acknowledgments

We would like to thank Jeanne-Marie Conquer and Hae Sun Kang, violinists from *l'Ensemble Intercontemporain* for their help in this study. We thank Alain Terrier for his contributions in developing the augmented violin prototypes. We also thank René Caussé and Norbert Schnell for support and Suzanne Winsberg for interesting discussions.

References

1. Anders Askenfelt. Measurement of the bowing parameters in violin playing. ii: Bow-bridge distance, dynamic range, and limits of bow force. *J. Acoust. Soc. Am.*, 86(2), 1989.
2. Ami Flammer and Gilles Tordjman. *Le Violon*. J.-C. Lattès and Salabert, 1988.
3. Emmanuel Fléty, Nicolas Leroy, Jean-Christophe Ravarini, and Frédéric Bevilacqua. Versatile sensor acquisition system utilizing network technology. In *Proceedings of the Conference on New Instruments for Musical Expression, NIME*, 2004.
4. Camille Goudeseune. *Composing with parameters for synthetic instruments*. PhD thesis, University of Illinois Urbana-Champaign, 2001.
5. Knut Guettler. On the creation of the helmholtz motion in bowed strings. *Acta Acustica*, 88, 2002.
6. Charles Nichols. The vbow: Development of a virtual violin bow haptic human-computer interface. In *Proceedings of the Conference on New Instruments for Musical Expression, NIME*, 2002.
7. Joseph A. Paradiso and Neil Gershenfeld. Musical applications of electric field sensing. *Computer Music Journal*, 21(2), 1997.

8. Chad Peiper, David Warden, and Guy Garnett. An interface for real-time classification of articulations produced by violin bowing. In *Proceedings of the Conference on New Instruments for Musical Expression, NIME*, 2003.
9. Cornelius Poepel. Synthesized strings for string players. In *Proceedings of the Conference on New Instruments for Musical Expression, NIME*, 2004.
10. Bernd Schoner, Chuck Cooper, Chris Douglas, and Neil Gershenfeld. Cluster-weighted sampling for synthesis and cross-synthesis of violin family instrument. In *Proceedings of the International Computer Music Conference, ICMC*, 2000.
11. Dan Trueman and Perry R. Cook. Bossa: The deconstructed violin reconstructed. In *Proceedings of the International Computer Music Conference, ICMC*, 1999.
12. M.M. Wanderley and M. Battier, editors. *Trends in Gestural Control of Music*. Ircam, 2000.
13. Diana Young. The hyperbow controller: Real-time dynamics measurement of violin performance. In *Proceedings of the Conference on New Instruments for Musical Expression, NIME*, Dublin, Ireland, 2002.
14. Diana Young and Stefania Serafin. Playability evaluation of a virtual bowed string instrument. In *Proceedings of the Conference on New Instruments for Musical Expression, NIME*, 2003.

ARTICLE II

Transients control of violin players: relationships between bow acceleration and string irregular vibrations.

Nicolas Rasamimanana,* Florian Kaiser, and Frédéric Bevilacqua

Ircam - CNRS UMR STMS

(Dated: December 12, 2007)

This paper reports an experimental multimodal study on violinist control during transients in actual playing situations. Bow acceleration and irregular string vibration are investigated during transients of various bowing techniques. Experiments were conducted with a violin teacher and seven students performing scales with bowing styles *Détaché* and *Martelé* at different dynamic levels. The string irregular vibrations are estimated from the recorded sound using High Resolution Methods. Simultaneously to the sound recording, the acceleration of the bow is measured using accelerometers placed at the frog. The results showed consistent shifts between time profiles related to bow acceleration and string irregular vibrations. In particular, these time shifts are found to be dependent on the bowing styles. The influences of bow direction and dynamic level are also examined and discussed.

I. INTRODUCTION

If playing music involves the production of the right tones at the right time, it also essentially deals with linking tones together. Musicians call *articulation* the manner of merging successive notes. On self-sustained music instruments, like winds or bowed strings, going from one tone to another implies going through a transient phase between two nearly periodic regimes. In the case of bowed strings, this transient phase is characterized by a non-periodic irregular motion of the string, where one Helmholtz motion stops and a new one develops. Through years of training, string players earn an in-depth control over this transient phase. Adjusting bow main parameters, expert players are able to vary the way it sounds: from smooth and light to harsh and crunchy. From this control, they therefore can produce different kinds of articulations.

Research in music has shown a vivid interest in transient parts, around the idea that they contain key expressive elements. Perception works demonstrated that transients play a preponderant role in instrument categorization and recognition^{12,16}. Audio signal synthesis techniques improved quality thanks to dedicated treatments on transient parts, especially techniques based on signal models^{18,19}. Acoustic studies investigated the origins of transient parts and their contribution to the acoustic signature of the instrument^{2,10}. Nevertheless, there is still very little studies on the gestural control of such transients. Using simulations and experiments on bowing machines, Guettler and Askenfelt studied non-periodic string motions with their associated combinations of constant bowing parameters^{13,14}. However, important questions on how musicians, directly or indirectly, control the sound of transient regimes remain. This paper tackles such a question by investigating both bow movements

and sound properties in actual playing situations with a particular attention given to the temporal behaviors of these multimodal components.

The non-periodic irregular string motion that occurs during the transient phase results in a broad band *noise*. For bowed string musicians, this typical noise is well-known. Violin pedagogue Ivan Galamian, talking about sound production on a violin, alludes to this noise saying that percussive sounds like consonants are necessary to shape the melody line formed by the vowel sounds⁹. Similar analogies between speech and music are also developed in music acoustics⁶. In particular, Joe Wolfe investigated clues supporting this analogy²⁰ and especially brought forward issues on timing and energy in broad spectral bands. These points are particularly insightful for the study of players' control on articulations in musical situations.

There are interesting parallels to draw between notions of transients in sounds and transition movements in music playing. For self-sustained instruments, the player's control is continuous, for example, bowed string players have to control the sound all the way from one stroke to another. Nevertheless, the movement for a transition between bow strokes is often considered by violinists as important as the movement for a stroke itself as indicated by violinist Ami Flammar⁷.

The general approach in this study is to consider sound and gesture parameters together in the context of bowing styles. Learning bowing styles indeed concerns both sound and gesture. On the one hand, the name of bowing styles often refer to an "action" (e.g. *Martelé*), on the other hand, the end goal is to achieve sound with specific characteristics. Bowing styles hence reflect "iconic", standard sounds with specific transient properties. However, it may requires years for student players to fully master a bowing technique and use it in a musical context. As a consequence, bowing styles offer a fertile ground and a structured basis for studies on string players' sound and gesture control.

*Nicolas.Rasamimanana@ircam.fr

II. AIM

The aim of this paper is to study the relationships between players' control, i.e. bow parameters, and the transient qualities determined by string irregular vibrations. The two classical bowing styles *Détaché* and *Martelé* are used as typical bowing strategies and investigated in actual playing situations. Complementarily to Guettler and Askenfelt studies^{1,13}, and similarly to Goebl's study on piano¹¹, this study aims at examining possible correspondance between dynamical and temporal aspects of players' bowing gesture and sound transients.

In Rasamimanana *et al.*¹⁷, bow acceleration was found to be a salient parameter to characterize the two bowing techniques *Détaché* and *Martelé*: differences and similarities between both techniques were shown with features derived from bow acceleration profiles. Moreover, bow acceleration is one of the essential parameters influencing string irregular vibrations as shown by Guettler¹³. In this study, temporal aspects of bow acceleration are therefore investigated with respect to string irregular vibrations.

III. MATERIAL AND METHOD

This study is based on the simultaneous recording of both violin sound and bow acceleration. The following sections describe the procedure, the material used to perform such measurements and the sound analysis that is performed to extract a signal related to sound transients.

A. Procedure

Seven violin students and their teacher from the "Conservatoire National de Région" of Dijon, France participated in the study. All of the students are advanced level, with between 9 and 14 years of practice. Each of them is recorded performing two exercices. First, they are asked to play two ascending and descending D major scales on one octave. First scale is performed in *Détaché* and second scale in *Martelé*, both at the dynamic level *forte*. *Détaché* is the most common technique. The sound is kept relatively constant during one stroke. The transition from one stroke to the other can be achieved with different degrees of smoothness / harshness but it generally remains smoother than in *Martelé*. In *Martelé*, the stroke is incisive and sounds almost percussive, hence the name. It is generally short, with a harsh beginning and ending. To reduce accelerometers possible bias due to gravity, subjects were asked to remain on the D string, therefore minimizing the angle variations of the bow. Second, they are asked to perform the same scales at the dynamic level *piano*. Both exercices are performed at the same tempo, 80bpm.



FIG. 1. Module placed at the frog of the violin bow to measure players' bowing movements. It consists of two accelerometers and a RF transmitter, powered with batteries.

B. Sound and gesture measurements

The system used to record players' bowing movements consists in two parts. First part is a module that measures bow acceleration with two accelerometers Analog Device ADXL202. This module is mounted at bow frog with a carbon clip. The placement of the two accelerometers is adjusted to measure bow dynamics along 3 directions: bow stick, strings and orthogonally to the stick. Accelerometers data are digitized on 16 bits every $3ms$ ($F_{sd} = 333Hz$) and sent wirelessly with a RF transmitter powered with batteries. This module is shown on Figure 1. The second part consists in a computer interface⁸ with a dedicated card receiving data from the RF transmitter. Data are sent over an ethernet connection to a laptop for recording using the Open Sound Control protocole. Accelerometer data is median filtered with a window of 8 samples at $F_{sd} = 333Hz$ to remove eventual acceleration peaks due to HF transmission errors. This system enables a non intrusive measurement of bow acceleration with an overweight of 17 grams at the frog. This system is similar to the one used in Bevilacqua *et al.*⁵.

A DPA 4021 microphone is clipped behind the violin bridge, enabling a recording of the violin direct sound. Sound is digitized at $F_{sa} = 44100Hz$ with a Maudio 410 firewire and sent to the laptop where it is recorded simultaneously to acceleration data using Max/MSP software (Cycling '74). A post synchronization is performed thanks to a synchronization signal obtained at the beginning of each measurement.

C. Sound analysis

Departing the irregular string vibration from the whole string vibration is not a trivial task. One approach in the

measurement of the quantity of irregular string vibration is to use signal processing techniques assuming a signal model with deterministic and stochastic components. In such procedures, the stochastic components contain the irregular vibrations of the string.

The separation between deterministic and stochastic components can be performed using analysis / synthesis techniques. The general procedure is to estimate the parameters of a signal model describing the deterministic components (analysis) and generate a new signal on the basis of this model (synthesis). Subtracting this modelled signal to the original signal gives a *residual* therefore containing the stochastic components. The quantity of irregular vibrations of the string along time can subsequently be estimated from the energy of the *residual*.

The model mostly used derives from Fourier spectral analysis where the deterministic components are represented as a sum of sinusoids with variable amplitudes, frequencies and phases. Because of the short time span of transient parts, between 50ms and 90ms in playing situations¹⁴, the chosen model in this paper is based on the formalism of High Resolution Methods (HRM). In HRM, the deterministic components are modelled as exponentially modulated sinusoids. This actually give HRM a higher frequency resolution than Fourier especially on short windows therefore enabling a more precise estimation of sinusoid parameters. The method applied in this paper is based on previous works on the use of High Resolution Methods in audio signal analysis^{3,15}, using ESPRIT for the estimation of the sinusoid parameters⁴.

1. Signal model

The deterministic components are modelled as a sum of exponentially modulated sinusoids. For all $t \in \mathbb{Z}$,

$$s(t) = \sum_{j=1}^K \alpha_k z_k^t \quad (1)$$

where $K \in \mathbb{N}^*$ is the order of the model. $\alpha_k \in \mathbb{C}^*$ are complex amplitudes. $z_k \in \mathbb{C}^*$ are distinct complex poles. z_k can be written as $z_k = e^{\delta_k} e^{2i\pi f_k}$, with $\delta_k \in \mathbb{R}$ the sinusoid modulating factor and $f_k \in \mathbb{R}$ the sinusoid frequency.

The observed audio signal is then represented as the combination of the deterministic component model $s(t)$ and an independent, centered, white gaussian noise $w(t)$ with variance σ^2 .

$$x(t) = s(t) + w(t) \quad (2)$$

2. Parameter estimation

In this paper, the estimation of the parameters, i.e. amplitudes and poles, is based on a property of the modelled signal covariance matrix $R_{ss}(t)$: the rank of $R_{ss}(t)$

is exactly K , the number of distinct poles, if it is of size $n > K$ and computed from $l > K$ observations. This has a direct consequence on the observed signal covariance matrix $R_{xx}(t) = R_{ss}(t) + \sigma^2 I_n$: a study on its rank permits to separate the observed signal space into two orthogonal subspaces, the *signal* space spanned by the exponentially modulated sinusoids and its orthogonal complementary, the *noise* space. Namely, the eigenvalues of the observed signal covariance matrix $R_{xx}(t)$ are

$$\lambda'_m = \begin{cases} \lambda_m + \sigma^2, \forall m \in \{0, 1, \dots, K-1\} \\ \sigma^2, \forall m \in \{K, \dots, n-1\} \end{cases}$$

where λ_m are the eigenvalues of the covariance matrix of $R_{ss}(t)$ sorted in decreasing order.

The poles are computed using the K first eigenvectors of $R_{xx}(t)$ combined with the property that the *signal* space is actually spanned by the poles. This is done with the ESPRIT algorithm⁴, based on the rotational invariance property of the *signal* space. The poles amplitudes are finally estimated with a least square regression.

3. Application to a violin recording

Previous studies showed that the ESPRIT algorithm provides an accurate estimation of the frequency of the deterministic components under the condition of an additive white noise³. To optimize the performance of the parameter estimation, the recorded audio signals are cut into eight frequency subbands of equal width. The analysis is then carried out independently on each subband, assuming a constant noise power on each of them.

The window size used to perform the analysis is 128 samples at $F_{sa} = 44100Hz$, i.e. 2ms. The number of exponentially modulated sinusoids K is usually unknown, although it plays a key role in the algorithm performances. For this study, K is set to 20 sinusoids per subbands. This value actually overestimates the theoretical value of 18 for D4 (294Hz), but ensures a correct estimation of the poles and their amplitudes¹⁵.

Once all the model parameters are estimated, in particular the sinus frequencies and amplitudes, a new signal is synthesized with an Overlap Add (OLA) technique, and the residual is obtained from the subtraction of the synthesized signal to the original signal. The energy is computed on a window of 1024 samples at $F_{sa} = 44100Hz$, i.e. 23ms, and then downsampled to the acceleration sampling frequency $F_{sd} = 333Hz$.

IV. RESULTS AND DISCUSSIONS

The goal of this study is to investigate relationships between bow acceleration and string irregular vibrations. First, typical temporal profiles of bow acceleration and residual energy are reported for single *Détaché* and *Martelé* bowings. Second, statistics are reported for scales

played by a professional player and seven students. After a general discussion, the results related to two specific cases are discussed separately: the differences between upbow and downbow and the case of the dynamic *piano*.

A. Temporal profile analysis

Recorded waveforms and bow acceleration are shown on Figure 2 for a series of strokes in *Détaché* (top) and *Martelé* (bottom). The residual energy computed as described previously is also plotted. From the figure, it can

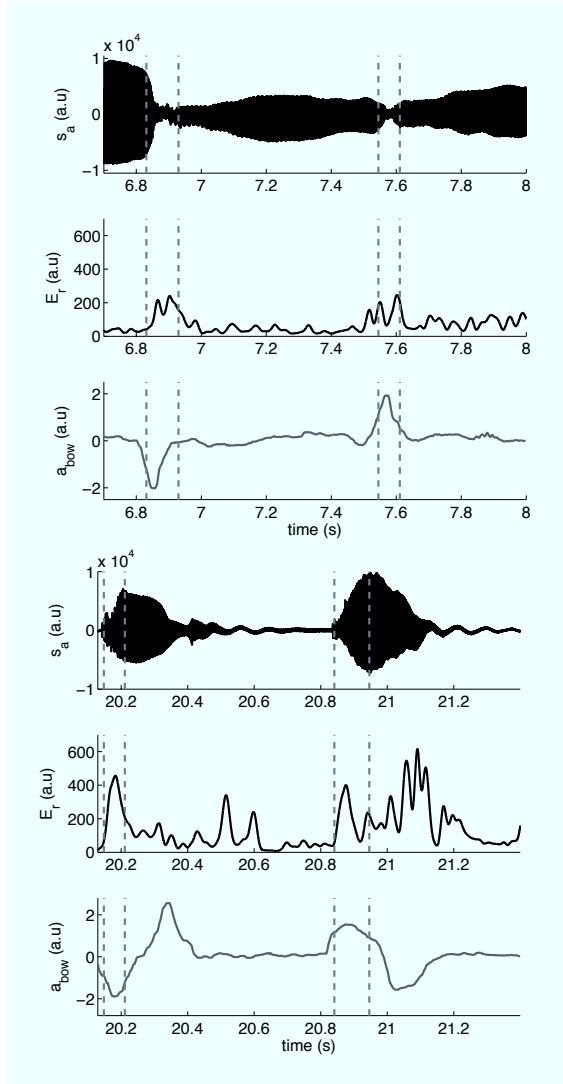


FIG. 2. From top to bottom: *Détaché* audio signal waveform, residual energy, bow acceleration, and *Martelé* audio signal waveform, residual energy, bow acceleration. Vertical bars delimit the analysis segments.

be noted that as expected in *Détaché* the residual energy

is maximum at transitions between strokes, which corresponds to transient phases occurring between periodic regimes. For *Martelé*, the residual energy is mainly located at the start and end of strokes, corresponding to the instants when the string periodic vibration is initiated and stopped. Moreover, as already noted in a previous study¹⁷, one bowing in *Détaché* is characterized by one acceleration peak, while two acceleration peaks (acceleration and deceleration) occur in one *Martelé* bowing.

The transients phases for the two bowing techniques are segmented to build a set of study data. Transitions between strokes are isolated for *Détaché* and only starting transients are kept for *Martelé*. The ending transients in *Martelé* are left aside of the study because the various performances were not consistent when stopping the string: bow leaving the string, bow jiggling before stopping or bow choking the string. The segmentation is done in two steps: first a manual segmentation is done to crop the transients, second an automatic process based on the amount of data within the segments adjusts the segments limits to have identically segmented data. Vertical dotted lines delimiting the analysis segments are shown Figure 2.

During transients, the residual energy and bow acceleration are characterized by different temporal distributions. In Figure 3 two examples of such distributions are shown for the bowing techniques *Détaché* and *Martelé*. Note that for clarity, the distributions are normalized (maximum is equaled to one). During transients both acceleration and residual energy display maxima. In the case of *Détaché* the maximum acceleration corresponds to the change in bow direction. In the case of *Martelé*, the maximum acceleration occurs during the attack. A striking feature is the time shift between the distributions of acceleration and sound residual. Notably, the shift seems to be dependent on the bowing style. To quantitatively appreciate this, first order moments $t_m = \frac{\sum s_n * t_n}{\sum s_n}$ are calculated, where s_n and t_n are respectively the digitized signal and the samples index. The first order moments for bow acceleration and the residual energy are not synchronous but present a specific time shift for each transients. The difference $\Delta t_m = t_m^{\text{residual}} - t_m^{\text{acceleration}}$ between the two moments is positive for *Détaché* and negative for *Martelé*. In the next sections these differences are discussed in the cases of different players, playing dynamics and bow directions.

B. Teacher

The difference between the first moments of bow acceleration and of the residual energy Δt_m is shown on Figure 4 (left) on the scale exercise for both bowing techniques *Détaché* and *Martelé* at the dynamic level *forte* for the teacher. For all 13 stroke transitions occurring in *Détaché*, this difference is positive, with a median value of 15ms, while for the 14 abrupt stroke attacks of *Martelé*, this difference is negative, with a median value of -18ms.

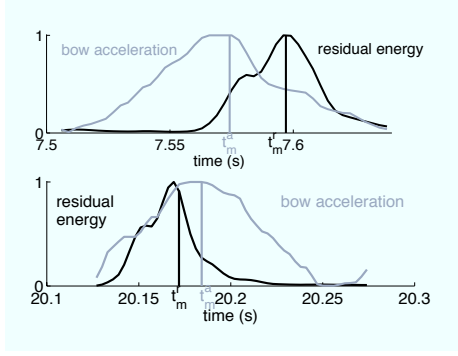


FIG. 3. Normalized temporal distributions of residual energy (dark) and acceleration absolute value (light) for one transient in *D  tach  * (top) and one transient in *Martel  * (bottom). t_m^r and t_m^a respectively designate first order moments for residual energy and bow acceleration.

The interquartile over the scale in *D  tach  * is $12.3ms$ and $10.8ms$ for the scale in *Martel  *. These values show that there is a statistically relevant timing difference between the two bowing techniques.

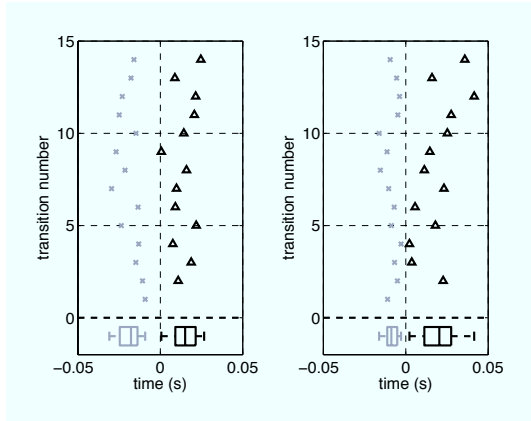


FIG. 4. Δt_m computed for the teacher transients on the scales in *D  tach  * (dark   ) and *Martel  * (light   ) for the dynamic levels *forte* (left plot) and *piano* (right plot). Each symbol correspond to a transient. Boxplots give synthetic views for each scale.

In spite of a softer sound, similar temporal relationships are found for the dynamic *piano* (Figure 4 right), namely, positive Δt_m for *D  tach  * and negative Δt_m for *Martel  *. The median and interquartile values for *D  tach  * are $16ms$ and $20ms$ and $-8ms$ and $6ms$, for *Martel  *. However, if the median and interquartile values are of the same order of magnitude in *forte* and *piano* for *D  tach  * they are smaller in *piano* for *Martel  *.

Two points emerge from these results. First, the introduced timing difference Δt_m is systematic for the two

bowing techniques, independently of the dynamic level. The timing Δt_m can therefore be used to differentiate the two bowing techniques *D  tach  * and *Martel  *. Second, unlike for *D  tach  *, different timings are found between *Martel   piano* and *Martel   forte*. This difference can be explained by the fact that *Martel  * usually results in a loud sound. Performing *Martel   piano* is less usual and is almost a different bowing technique: if the strokes remain incisive, the attack transients are globally softer than in "traditional" *Martel  *. This difference consistently appears on the timing values Δt_m : lower negative values correspond to the harsh transients of *Martel   forte* and higher negative values to the softer transients of *Martel   piano*. The performance of dynamics *piano* is specifically discussed in the dedicated section.

C. Students

The temporal relationships described previously are now studied for the seven violin students. Figure 5 (top) shows the different timings Δt_m on the whole scale exercise for each bowing technique at the dynamic level *forte*. First, as for the teacher, and in spite of idiosyncrasies, all of them did perform with a positive timing for *D  tach  * and negative timing for *Martel  *. When considering the totality of the students, the median values for *D  tach  * and *Martel  * are indeed $19ms$ and $-20ms$. Moreover, for *D  tach  * the mean interquartile value over the students is similar to the teacher's, i.e. $15ms$. However, the students show more variability on the exercise than the teacher on *Martel  * bowing: the mean interquartile value is $21ms$ instead of $11ms$ for the teacher, i.e. twice as much variability.

Because the bow must be more carefully controlled at the beginning and at the end of each stroke, *Martel  * is actually more difficult a bowing technique than *D  tach  *. The difference in variability found in *Martel  * can be the result of more different executions for less skilled players. Besides, student3, one of the less skilled student, shows the most variability in *Martel  *, i.e. $39ms$, therefore suggesting a link between the larger variability found in students and their control abilities on the string vibration. Compared to the students, the teacher statistically shows a stronger consistency in the execution of the two bowing techniques.

D. Asymmetry between up and down bows

Symmetrical bowing between up and down bows is a well-known issue in bowed string instruments. Part of the greater variability found in students can actually be explained with their asymmetric control between up and down bows. This is indeed manifest when looking at the timings separately, as shown on Figure 6 for student1 *forte*. Interestingly the boxplots overlap for *D  tach  * while they are separated for *Martel  *. For *Martel  * down-

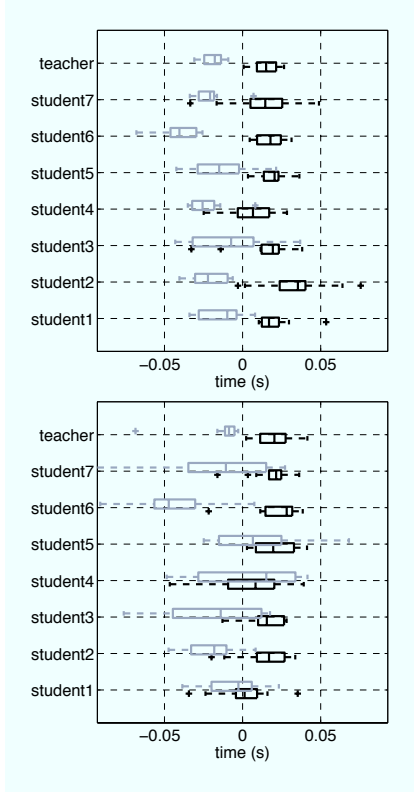


FIG. 5. Δt_m for the teacher and the 7 students for each scale in *Détaché* (dark) and *Martelé* (light). Each boxplot represents one scale. (top): dynamics level *forte*, (bottom): dynamic level *piano*.

bows, the Δt_m value is $-28ms$, i.e. of the same magnitude as the teacher's for the same dynamic level. However, the median value for *Martelé* upbows is only $-4ms$, hence similar to the teacher's value for the dynamic level *piano*.

Moreover, less variability is found when separating upbows and downbows. The interquartiles are indeed reduced respectively to $8ms$ and $10ms$. Therefore, the upbows are all performed a similar manner, and it is likewise for the downbows. However, the upbows and downbows are performed differently from each other.

E. Piano dynamic

As already found for the teacher, the analysis of the students at the dynamic level *piano* reveals a more varied situation, especially in *Martelé* as shown on Figure 5 (bottom). If the median values for the whole students are still consistent with previous analyses, i.e. $16ms$ for *Détaché* and $-10ms$ for *Martelé*, the timings show very large variability, as indicated by the interquartiles: between $7ms$ and $30ms$ for *Détaché* and between $23ms$ and $62ms$ for *Martelé*.

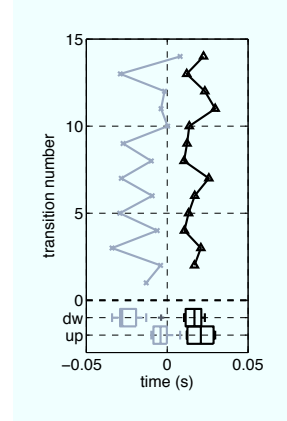


FIG. 6. Asymmetry between alternating up and down bows in *Détaché* (dark) and *Martelé* (light). Separate boxplots are computed for upbows (up) and downbows (dw).

Nevertheless, this can be the expression of the fact that performing *Martelé piano* is challenging: the sought tone must both have a harsh attack and a soft sound. For this exercise, the teacher shows few variability, therefore performing in a stereotypical way and proving a strong consistent control on the string vibrations. The students, not as experts, show difficulties as indicated by the large interquartile values. Their control on the string vibrations is not homogeneous, sometimes performing transients closer to *Détaché*, sometimes closer to *Martelé*. This result actually complements those in reference¹⁷, where similarities were found between *Détaché piano* and *Martelé piano* based on features derived from bow acceleration.

V. CONCLUSIONS

This paper investigates bowed string players' control over the string irregular vibrations during transients in live playing conditions. Such control is actually essential for players to *articulate* tones with each others. This control is studied under different situations where transients are produced from the execution of *Détaché* and *Martelé* at two dynamic levels *forte* and *piano*.

From the analysis of a class of advanced violin players performing scales, it is shown that consistent time control is achieved between bow acceleration and the quantity of string irregular vibrations expressed by the residual energy. Salient, specific temporal relationships are found for *Détaché* stroke transitions and *Martelé* stroke attacks: for soft transients such as found in *Détaché* the maximum of bow acceleration occurs before the maximum quantity of string irregular vibrations, while it is the reverse for harsh transients such as produced in *Martelé*. These relationships are independent of the dynamic level, *forte* or *piano*, although more variability was found

in the latter.

More variability is also found in the execution of *Martelé* between the students for a given dynamic level and for each student between up and down bows. More generally, the analyses of *Martelé* transients show more variability than *Détaché*, therefore expressing the greater difficulty for student players to perform this bowing technique.

The different time shifts found between the residual energy and bow acceleration indicate, as expected, that these parameters are not directly correlated to each other. The actual relationship between the residual energy and bow acceleration is complex and influenced by several other bowing parameters notably bow force¹³. Therefore, it is remarkable to find systematic timing relationships between the residual energy and bow acceleration in playing situations where parameters such as bow force are likely to vary importantly.

In a previous study, *Détaché* and *Martelé* strokes were characterized from the intensity of bow acceleration only¹⁷. The presented results complement this study by actually adding timing considerations between gesture and sound. Moreover, the results show that the distinction between *Détaché* and *Martelé* can be performed from the very beginning of the stroke. This study brings new insights that might be useful for designing control of sound synthesis, and brings interesting questions on applications for music pedagogy.

VI. ACKNOWLEDGEMENTS

The authors would like to acknowledge Roland Badeau for his help on High Resolution Methods. The authors also thank Norbert Schnell and René Caussé for discussions and support. The authors particularly thank Anne Mercier, Florence Baschet and the students of the Dijon Music Conservatory for their participation in this study.

- ¹ Askenfelt, A. (1989). "Measurement of the bowing parameters in violin playing. ii: bow-bridge distance, dynamic range, and limits of bow force", *Journal of the Acoustical Society of America* **86**.
- ² Askenfelt, A. (1993). "Observations on the transient components of the piano tone", in *Proceedings of the Stockholm Music Acoustics Conference (SMAC93)*, volume 79, 297–301.
- ³ Badeau, R. (2005). "Méthodes à haute résolution pour l'estimation et le suivi de sinusoïdes modulées. application aux signaux de musique", Ph.D. thesis, Ecole Nationale Supérieure des Télécommunications.
- ⁴ Badeau, R., Richard, G., and David, B. (2005). "Fast adaptive esprit algorithm", in *IEEE Workshop on Statistical Signal Processing SSP'05* (Bordeaux, France).
- ⁵ Bevilacqua, F., Rasamimanana, N. H., Fléty, E., Lemouton, S., and Baschet, F. (2006). "The augmented violin

- project: research, composition and performance report.", in *Proceedings of the Conference on New Interfaces for Musical Expression, NIME*.
- ⁶ Chafe, C. (1990). "Pulsed noise in self-sustained oscillations of musical instruments", in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, volume 2, 1157–1160.
 - ⁷ Flammer, A. and Tordjman, G. (1988). *Le violon* (J.C. Lattès, Salabert).
 - ⁸ Fléty, E., Leroy, N., Ravarini, J.-C., and Bevilacqua, F. (2004). "Versatile sensor acquisition system utilizing network technology", in *Proceedings of the Conference on New Interfaces for Musical Expression, NIME*.
 - ⁹ Galamian, I. (1999). *Principles of Violin Playing and Teaching* (Shar Products Co).
 - ¹⁰ Goebel, W., Bresin, R., and Galembo, A. (2004). "Once again: The perception of piano touch and tone. can touch audibly change piano sound independently of intensity?", in *Proceedings of the International Symposium on Musical Acoustics*.
 - ¹¹ Goebel, W., Bresin, R., and Galembo, A. (2005). "Touch and temporal behavior of grand piano actions", *Journal of the Acoustical Society of America* **118**, 1154–1165.
 - ¹² Grey, J. M. (1975). "An exploration of musical timbre using computer-based techniques for analysis, synthesis and perceptual scaling", Ph.D. thesis, Stanford University.
 - ¹³ Guettler, K. (2004). "Looking at starting transients and tone coloring of the bowed string", in *Proceedings of Frontiers of Research on Speech and Music*.
 - ¹⁴ Guettler, K. and Askenfelt, A. (1997). "Acceptance limits for the duration of pre-helmholtz transients in bowed string attacks", *Journal of the Acoustical Society of America* **101**, 2903–2913.
 - ¹⁵ Laroche, J. (1989). "A new analysis/synthesis system of musical signals using prony's method-application to heavily damped percussive sounds", in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, volume 3, 2053–2056.
 - ¹⁶ McAdams, S., Winsberg, S., de Soete, G., and Krimphoff, J. (1995). "Perceptual scaling of synthesized musical timbres: common dimensions, specificities and latent subject classes", *Psychological Research* **58**, 177192.
 - ¹⁷ Rasamimanana, N. H., Fléty, E., and Bevilacqua, F. (2006). "Gesture analysis of violin bow strokes", in *Lecture Notes in Artificial Intelligence, LNAI 3881*, edited by S. Gibet, N. Courty, and J.-F. Kamp, 145–155 (Springer Verlag).
 - ¹⁸ Serra, X. and Smith III, J. O. (1990). "Spectral modeling synthesis: A sound analysis/synthesis system based on a deterministic plus stochastic decomposition", *Computer Music Journal* **14**, 12–24.
 - ¹⁹ Verma, T., Levine, S., and Meng, T. (1997). "Transient modeling synthesis: a flexible analysis/synthesis tool for transient signals", in *Proceedings of the International Computer Music Conference, Greece*.
 - ²⁰ Wolfe, J. (2002). "Speech and music, acoustics and coding, and what music might be 'for'", in *Proceedings of the 7th International Conference on Music Perception and Cognition, Sydney*.

ARTICLE III

String Bowing Gestures at Varying Bow Stroke Frequencies: A Case Study

Nicolas Rasamimanana¹, Delphine Bernardin², Marcelo Wanderley², and
Frédéric Bevilacqua¹

¹ IRCAM, CNRS - UMR 9912, 1 Place Igor Stravinsky, 75004 PARIS, France

{Nicolas.Rasamimanana, Frederic.Bevilacqua}@ircam.fr

² IDMIT/CIRMMT, McGill University, Montreal, QC, Canada

{Delphine.Bernardin, Marcelo.Wanderley}@mcgill.ca

Abstract. The understanding of different bowing strategies can provide key concepts for the modelling of music performance. We report here an exploratory study of bowing gestures for a viola player and a violin player in the case of bow strokes performed at different frequencies. Bow and arm movements as well as bow pressure on strings were measured respectively with a 3D optical motion capture system and a custom pressure sensor. While increasing bow stroke frequency, defined as the inverse time between two strokes, players did use different bowing movements as indicated from the measurement of bow velocity and arm joint angles. First, bow velocity profiles abruptly shift from a rectangle shape to a sinus shape. Second, while bow velocity is sinusoidal, an additional change is observed: the wrist and elbow relative phase shifts from out-of-phase to in-phase at the highest frequencies, indicating a possible change in the players coordinative pattern. We finally discuss the fact that only small differences are found in the sound while significant changes occur in the velocity / acceleration profiles.

1 Introduction

The understanding of different bowing strategies can provide key concepts for the modelling of music performance. Such model can be applied in music pedagogy [1] or in the design of novel musical interfaces [2]. We previously reported the study of three standard bowing techniques in violin playing [3]. In particular, we discussed issues on gesture "continuity". This concept relates to the fact that an expert violinist is able to play several and subtle variations between two bowing techniques, e.g. *Détaché* and *Martelé*. We showed in reference [3] that such subtle variations can be directly tracked in the bowing dynamics and described with features derived from bow acceleration profiles.

In this paper, we report complementary results on violin playing by studying bowing gestures, i.e. movements directly involved in sound production, at different stroke frequencies and further question the concept of gesture "continuity". Slow bowing generally requires a relaxed right arm. Nevertheless, it is usually recognized among bowed string players that rapid, repeated bow strokes can require

the right arm to be tensed up. Such strategy is usually used to achieve a given rhythm, for example four sixteenths, or to perform a *tremolo*. From the players' viewpoint, these are different arm movements to perform cyclic, repetitive bowings. We here test this hypothesis with the study of an *accelerando/decelerando*, i.e. bow strokes performed with an increasing/decreasing frequency, and investigate on the continuity between slow and fast bowing.

The paper is structured as follows. First we describe related works and our experimental method based on optical 3D motion capture technology combined with a bow pressure sensor. Second, we present and discuss results obtained by measuring arm and bow movements of two instrumentalists. Third, we investigate sound characteristics at different bow stroke frequencies. Finally, we present conclusions and directions for future work.

2 Related works

When performing a *glissando*, it may happen that singers produce a discontinuity in pitch. It corresponds to the transition from one laryngeal mechanism to another, such as described in reference [4]. This drastic change in voice production enables singers to achieve the highest frequencies of the *glissando*. We hypothesize that a similar phenomenon occurs for bowed string instruments: players can change their bowing gestures to perform high bow stroke frequencies.

This configuration change in voice production relates to a well-known notion in motor control: the reorganisation of coordinative patterns, as occurring in gait shifting. In particular, Kelso studied the abrupt phase transitions in human hand movements according to the cycling frequency [5, 6]: for example, a periodic out-of-phase movement of human's fingers, i.e. one finger up while the other is down, shifts to an in-phase movement, i.e. both fingers up or down at the same time, when increasing frequency. These results are especially insightful for our study as we are dealing with a cyclic movement with increasing frequency (*accelerando*), and involving the upper arm, the forearm and the hand.

Several studies report on the movement analysis of instrument performance [7, 8], [9], [10] and in particular bowed string players [11]. Winold et al. [12] first studied coordination issues in bow arm movements in a musical context. They analyzed cellists' bowing coordinations while performing fragments by Brahms and Schubert at different tempi. They concluded that increasing tempi produced a proportional scaling of stroke amplitudes and durations. Nevertheless, they did not notice any change in within-limb coordinations. More recently, Baader et al. [13] studied coordination between fingering and bowing and showed anticipatory movements between the left hand and the right arm. While our approach is similar to Winold's, we here focus on bow, elbow and wrist movements on a simpler musical task: *accelerando/decelerando* on one single note.

3 Method

We used a Vicon System 460 optical motion capture system to measure the arm and bow movements. Six M2 cameras were placed around the instrumentalists providing a spatial resolution below 1mm with a frame rate of 500Hz. Markers were placed on the players' upper body, following the standard marker placement in the Vicon Plug-in Gait [14]. Six markers were placed on the instrument, four for the table and two for the strings. Three markers were placed on the bow. With this setup, the movement of the bow can be computed relatively to the instrument. The position of the contact point between the bow and the strings is calculated and is used as the center of an instrument-based frame of reference.

Bow pressure on string was measured with a custom sensor designed at Ircam [15], with a $\pm 3\%$ error. Bow pressure data were recorded separately from motion capture data. To ensure a post-recording synchronization between both sets of data, the sound was recorded simultaneously with each sensing system: we use the arg-maximum of the cross-correlation between the audio signal envelopes to align both sets of data.

Two students from McGill Schulich School of Music were recorded playing an *accelerando*, from medium paced *Détaché* to a "as fast as possible", tied with a *decelerando* back to medium paced *Détaché*. Musicians were asked to stay on a fixed note. One McGill student played the violin, the other student played the viola. Both McGill students were advanced level with more than ten years of instrument practice.

4 Results and Discussion

The presentation of results is structured as follows: first, we study bow movement and focus on a change in bowing gestures found on the velocity and acceleration profiles during an *accelerando/decelerando*. Second, we present the results relative to the arm joint angles providing additional insights on this change. Last, we show results of the sound characteristics at the moment of the bowing change.

4.1 Change in bowing gestures

Bow movements: Figure 1 shows the position, velocity and acceleration of one point of the bow, for the viola player. For clarity, the bow stroke frequency is also plotted, showing the *accelerando/decelerando*. Bow stroke frequency is defined as the inverse of a stroke duration and is computed as the inverse of the time separating two successive zero crossings of the bow velocity curve. Due to the physics of bowed strings, the amplitude variations of velocity and acceleration are bounded within an interval that guarantees the production of an acceptable sound [16]. There is no such constrain on the distance, and thus tempo variation is achieved by reducing the length of bow, as already noticed in reference [12]. Biomechanics also imposes this reduction: the combination fast tempo - long strokes is indeed very difficult to achieve. Figure 1 shows that the

absolute value of the acceleration amplitude remains relatively constant, while the observed dramatic decrease of the position amplitude is directly correlated with the increase of bow stroke frequency.

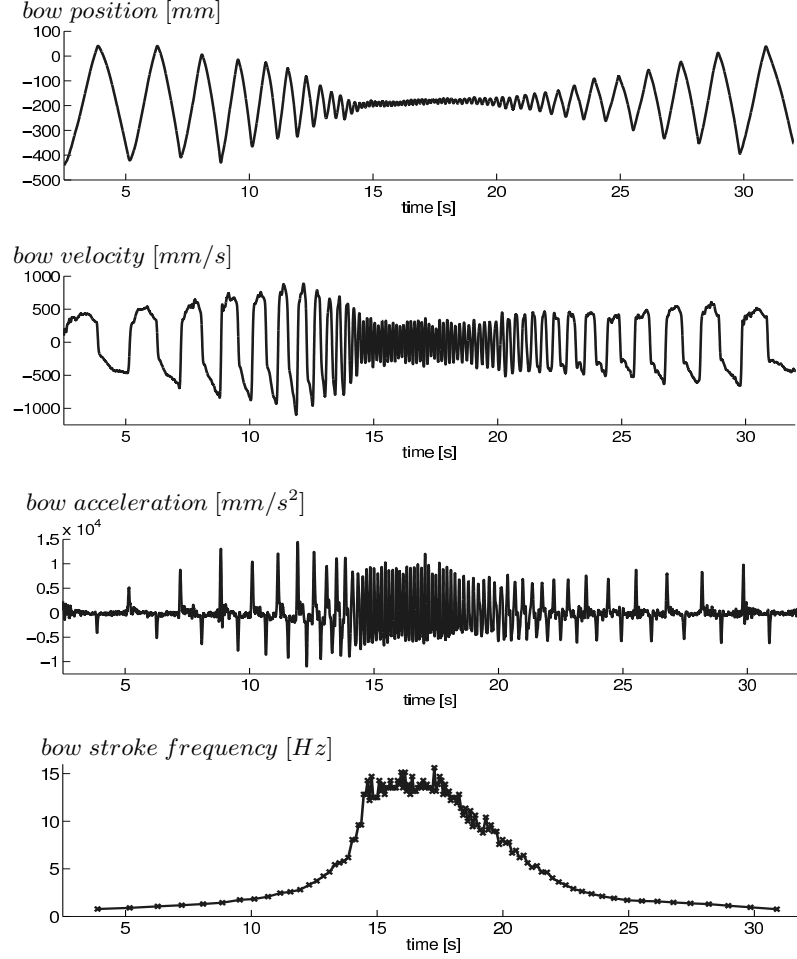


Fig. 1. From top to bottom: bow position, velocity, acceleration and bow stroke frequency for the *accelerando* / *decelerando* performed by the viola player.

However, a zoomed view of bow velocity and bow acceleration reveals an interesting profile change in the middle part of the *accelerando* / *decelerando*. Figure 2 shows the two players' bow dynamics at this moment. The profile change is observable for both players. Two different patterns are clearly visible on the graphs: the acceleration profile drastically changes becoming smoother at times

$t = 14s$ for the viola and $t = 11s$ for the violin. For both players, the reverse change occurs during the *decelerando*.

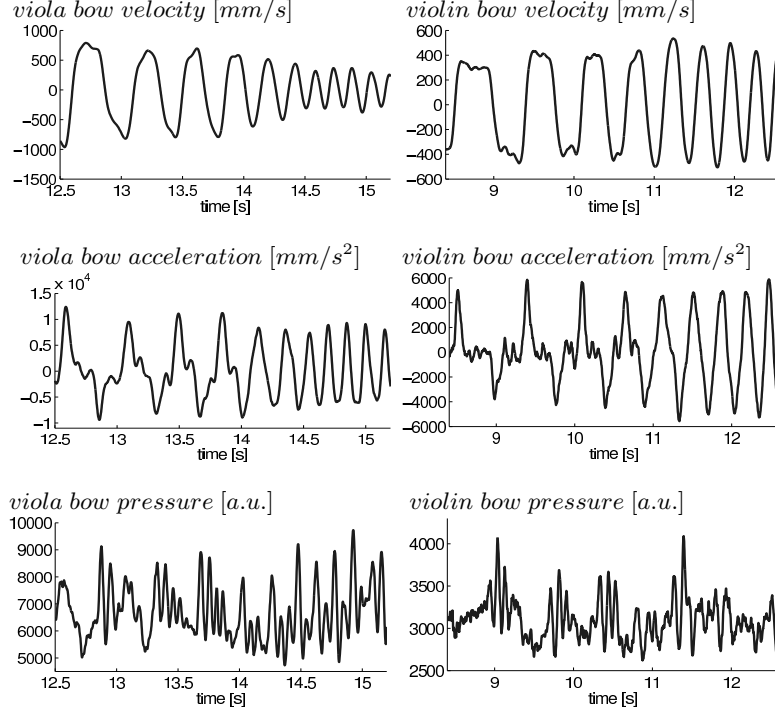


Fig. 2. From top to bottom: bow velocity, bow acceleration and bow pressure on strings. For both instrumentalists, a clear change in profile occurs for bow velocity and acceleration but not for bow pressure.

These two patterns allow us to section the *accelerando* / *decelerando* in three parts. In the first and third parts, denoted by $C1$ and $C3$, the velocity is close to a square signal, the acceleration profile is characterized by well defined short positive and negative peaks. Between these peaks, other smaller peaks are observable. In the second part, denoted by $C2$, the velocity and acceleration profiles are smooth and close to sinusoidal. It is worth to notice that the change in bowing gesture is abrupt: no intermediate shape can be observed in the acceleration profile in Figure 2. It is also interesting to observe that no obvious, concomitant change occurs in the profile of bow pressure.

Profile characterization: To quantitatively characterize this profile change, we perform a sinus non-linear fit on bow velocity. Bow velocity is segmented in individual strokes. Each segment is resampled to a fixed number of points:

we chose 600 points, i.e. the length of the longest stroke in the measurements. Each segment is fitted to half period of a sinus fitting function, that allows for variable amplitude and phase, but with a fixed frequency. The estimation of the amplitude and phase parameters is performed with a non-linear least square regression. The mean square *fit error* therefore provides a measurement of the profile change.

Figure 3 shows the *fit error* parameter for the whole *accelerando* and *decelerando*. We can see a minimum plateau in the middle part for both players. Also, the value of the *error* is drastically higher at the beginning and at the end. This illustrates the change in profile observed in the previous section.

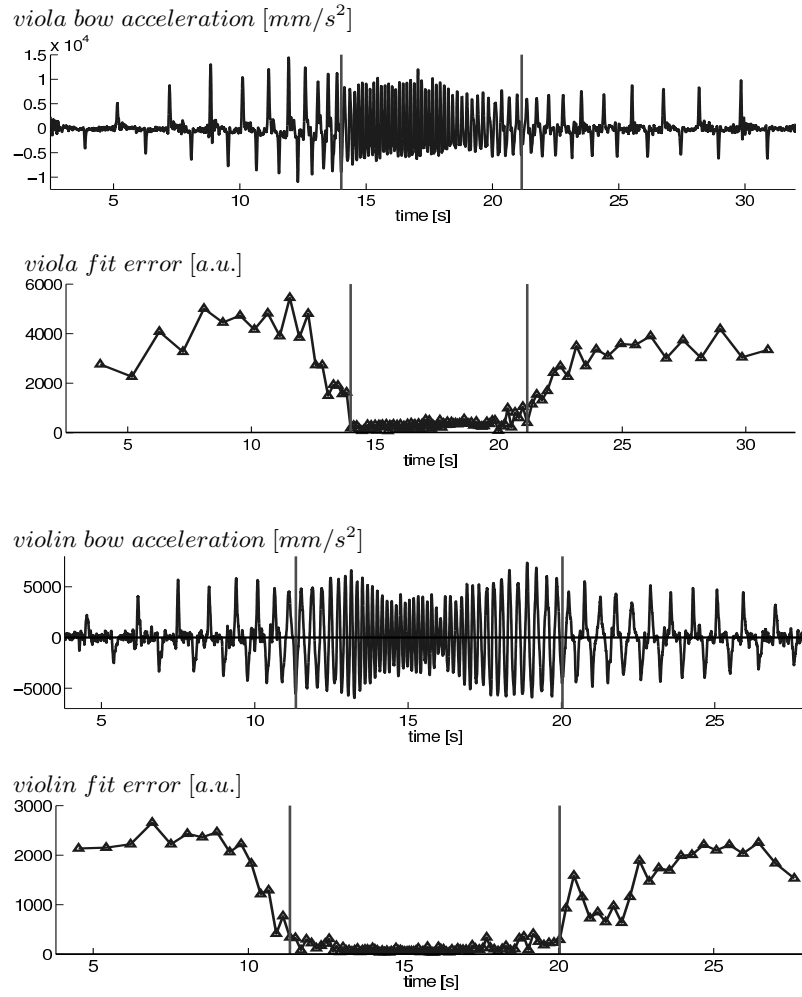


Fig. 3. Bow acceleration and the sinusoid *fit error*. The vertical bars were computed with a threshold based on *fit error* values.

In these figures, we can delimitate the $C1$, $C2$ and $C3$ parts by applying a threshold on the *fit error* parameter as shown by vertical bars in Figure 3. The profile changes in bow velocity and acceleration as seen in Figure 2 correspond approximately to a threshold of $560[a.u.]$ (value determined empirically).

Figure 4 plots the *fit error* along with the strokes period, defined as the time interval between up and down bows. For both players, we can see that while the period decreases and increases in a linear way, the *fit error* abruptly shifts from high to minimum values. This endorses our previous observation of abrupt transitions between the three parts $C1$, $C2$ and $C3$.

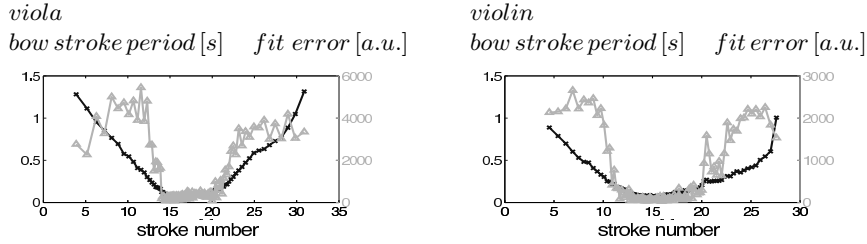


Fig. 4. In light color, the *fit error* and in dark color, the bow stroke period. While the period decreases and increases in a linear way, the *fit error* abruptly shifts from high to minimum values

We can also observe an hysteresis between the *accelerando* and the *decelerando*: the transition $C2 \rightarrow C3$ takes more time than $C1 \rightarrow C2$ for the two players. Moreover, we can notice that the frequency at which velocity profiles change is different between the two players. For the violin player, the $C2$ part comes early in the *accelerando*, at $f_{shift:C1 \rightarrow C2} = 7Hz$, and the bowing frequency keeps on increasing up to $f_{max} = 15Hz$. On the contrary, for the viola player the $C2$ part coincides with the *accelerando* climax, at $f_{shift:C1 \rightarrow C2} = f_{max} = 14Hz$. We cannot at this point know whether this difference is a function of the player or the instrument. More violin and viola players need to be considered to determine the typical frequencies for this profile change and their dependence on the instrument and the players' expertise.

Arm angles: Bow movements result from the coordination of the upper arm, forearm and hand. The analysis of arm joint angles can therefore give further insights to the observed change in bowing gesture, described in the previous section. From the motion capture data, we computed the elbow angle, i.e. the angle between the upper arm and the forearm segments, and the wrist angle, i.e. the angle formed by the forearm and the hand. We therefore consider the wrist and elbow angles, main contributors to the bowing movement, in a periodic flexion-extension movement.

The angle derivatives provide interesting information. Figure 5 plots the second derivative of the elbow angle. First, contrary to bow movement, arm movement shows major differences between the two players, as shown on Figure 5. This is explained by the large number of degrees of freedom in the arm. These differences actually express the players' personal bowing technique. However, for both players, we can notice changes at the transition time we determined from bow movements ($C1 \rightarrow C2$ and $C2 \rightarrow C3$). In Figure 5, we can see that similarly to bow acceleration, the profile of the second derivative of the elbow angle shows a dramatic change for both players ($t = 14s$ and $t = 11s$): it becomes smoother and almost sinusoidal.

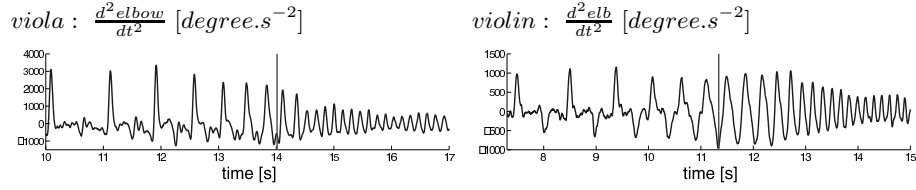


Fig. 5. Second derivative of the elbow angle. The vertical lines indicate the transition $C1 \rightarrow C2$. Similarly to bow acceleration, the profile of the second derivative of the elbow angle shows a dramatic change for both players.

Moreover, three additional parts can be identified in the arm movement from the analysis of joints' relative phases. The first derivatives for the wrist and elbow angles are plotted on Figure 6. Indeed, during the $C2$ part previously defined, we can see another clear change occurring at the climax of the *accelerando*: for the violin player, the wrist and elbow curves first are in anti-phase at $t = 12s$, suddenly shift to in-phase at $t = 13.2s$ and gradually shift back to anti-phase at $t = 19.2s$. This change occurs when the bowing frequency is the highest. A similar but less obvious change also occurs for the viola (changes at $t = 18.8s$ and $t = 22.2s$). For the viola player, the in-phase / anti-phase transition occurs in the reverse way, i.e. first in-phase, then anti-phase.

These phase transitions indicate a possible reorganisation of within-limb coordinations such as described in [5], under the influence of increasing bowing frequency. Additional data from more players will bring a further characterization of this observation. It is also interesting to note that this change in joints' relative phase does not have a clear influence on the dynamics of the bow.

4.2 Audio comparison of the two bowing gestures

In reference [3], we reported the relationships between bow acceleration curves and bowing techniques that correspond to specific sound characteristics. We particularly stressed the gesture-sound continuity between the different bowing techniques. In the previous sections, we identified different bowing gestures for

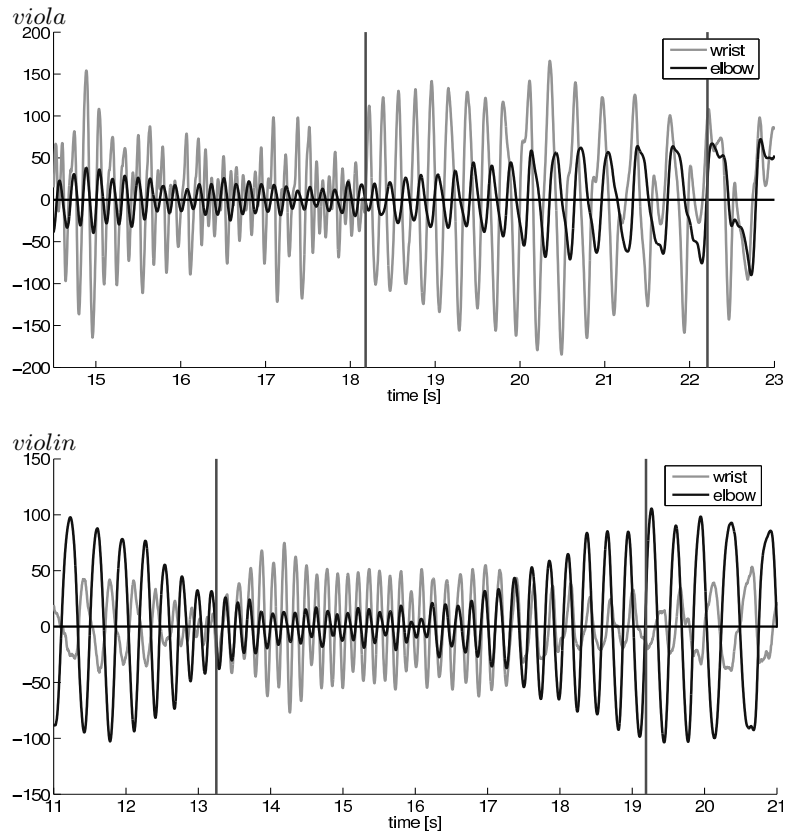


Fig. 6. First derivatives of elbow and wrist angles. Vertical bars indicate changes in the joints' relative phase.

slow and fast bowing with a brutal change between the two. We now investigate the effect of this change on the produced sound.

Interestingly, very few differences can be heard in spite of the drastic change previously described. This observation is supported by an audio spectrum comparison: for both players, there are indeed small differences in the spectra at the transitions ($C1 \rightarrow C2$ and $C2 \rightarrow C3$), as shown on Figure 7 for the viola player. We can see on top the mean audio spectrum $S1$ over the three strokes before

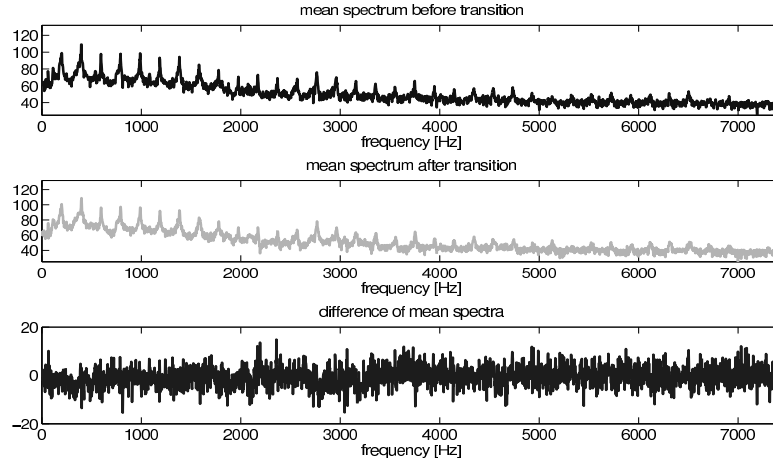


Fig. 7. Audio spectrum comparison for the viola player. Top: before $C1 \rightarrow C2$, Middle: after $C1 \rightarrow C2$, Bottom: spectrum difference

transition $C1 \rightarrow C2$, in the middle the mean audio spectrum $S2$ over the three strokes after the same transition, and at the bottom the difference between the two spectra ($S1 - S2$). We can graphically see that the two spectra have similar peaks in frequencies and amplitudes. This is confirmed with the difference between the two spectra: mean of $1dB$ with a standard deviation of $4dB$. However, it is worth to note that the difference is not uniform among the frequencies but slightly more important in the medium range, i.e. between $100Hz$ and $600Hz$. The spectral peaks having similar amplitudes, this is due to a difference in the noise level, which is lower in $C1$ than in $C2$. The origin of this last point will be further investigated with the help of physical models of bowed strings.

Further studies will also help to determine whether the sound similarities are actually due to an active, conscious or unconscious control of the musicians.

5 Conclusion and future directions

We present in this paper a study on the use of different bowing gestures by a violin player and a viola player to achieve different bow stroke frequencies. From the analysis of bow movement in an *accelerando/decelerando*, we showed the existence of two profiles in bow velocity and acceleration, therefore defining two bowing gestures. With a profile characterization based on a sinus non-linear fit, we noticed that the transition from one gesture to the other is abrupt with an hysteresis effect.

The study of arm joint angles also indicated the possible existence of a within-limb change of coordination for very fast bow strokes (15 Hz). In the case of the violin player, the elbow and the wrist first start in out-of-phase and shift to in-phase to achieve the fastest part of the *accelerando/decelerando*.

In the recorded performances of *accelerando/decelerando*, we could therefore clearly identify four parts in the players' movements: a square-shaped bow velocity, a sinus-shaped velocity with two possible arm coordinations, and a square-shaped bow velocity. Further studies must be performed to clarify the generalities of our findings.

Besides, an audio spectrum analysis does not reveal a clear concomitant change to the drastic change in bow velocity profiles. This might be due to the players being sufficiently experienced to smooth out the effects of changing bowing strategy. To test this hypothesis, new experiments with students of various levels, including beginners must be carried out. We can also hypothesize that the change in bowing gesture has an effect on finer timbre aspects like e.g. transitions between notes. These non obvious correspondances between gesture and sound and especially their evolutive aspect open interesting questions for the control of electronic sounds.

6 Acknowledgements

Nicolas Rasamimanana gratefully thanks funds from the Centre de Collaboration Interuniversitaire Franco Québécois. Thanks to Matthias Demoucron for his help and for lending us the pressure sensor he developed. We thank Florence Baschet, Christophe Desjardins, Serge Lemouton, Norbert Schnell and René Caussé for contributions, discussions and support.

This work was partly founded by the european project i-Maestro IST-026883. We also acknowledge the support of the project CONSONNES funded by the ANR (Agence Nationale de la Recherche, France).

The second and third authors are partially funded by the Natural Sciences and Engineering Council of Canada (Discovery and Special Research Opportunity grants) and by the Ministry of Economic Development of Quebec (Programme de soutien aux initiatives internationales de recherche et d'innovation). Thanks to support from the 6th Framework Project ENACTIVE Interfaces.

References

1. I-MAESTRO: <http://www.i-maestro.org>
2. Bevilacqua, F., Rasamimanana, N.H., Fléty, E., Lemouton, S., Baschet, F.: The augmented violin project: research, composition and performance report. In: NIME Proceedings. (2006) 402–406
3. Rasamimanana, N.H., Flety, E., Bevilacqua, F.: Gesture analysis of violin bow strokes. Lecture Notes in Artificial Intelligence, LNAI 3881 (2006) 145–155
4. Henrich, N.: Mirroring the voice from garcia to the present day: Some insights into singing voice registers. Logopedics Phoniatrics Vocology **31** (January 2006) 3–14(12)
5. Haken, H., Kelso, J.A.S., Bunz, H.: A theoretical model of phase transitions in human hand movements. Biological cybernetics **51**(5) (1985) 347–356
6. Fuchs, A., Jirsa, V.K., Haken, H., Kelso, J.A.S.: Extending the hkb model of coordinated movement to oscillators with different eigenfrequencies. Biological cybernetics **74**(1) (1996) 21–30
7. Palmer, C.: Movement amplitude and tempo change in piano performance. Acoustical Society of America Journal **115** (2001) 2590–2590
8. Dalla Bella, S., Palmer, C.: Tempo and dynamics in piano performance: the role of movement amplitude. In Lipscomb, S.D., Ashley, R., Gjerdingen, R., Webster, P., eds.: Proceedings of the International Conference on Music Perception and Cognition. (2004) 256–257
9. Wanderley, M.M., Vines, B.W., Middleton, N., McKay, C., Hatch, W.: The musical significance of clarinetists' ancillary gestures: An exploration of the field. Journal of New Music Research **34**(1) (March 2005) 97–113
10. Dahl, S.: Playing the accent - comparing striking velocity and timing in an ostinato rhythm performed by four drummers. Acta Acustica united with Acustica **90**(4) (2004) 762–776
11. Shan, G.B., Visentin, P.: A quantitative three-dimensional analysis of arm kinematics in violin performance. Medical Problems of Performing Artists **18**(1) (2003) 3–10
12. Winold, H., Thelen, E., Ulrich, B.D.: Coordination and control in the bow arm movements of highly skilled cellists. Ecological Psychology **6**(1) (1994) 1–31
13. Baader, A.P., Kazennikov, O., Wiesendanger, M.: Coordination of bowing and fingering in violin playing. Cognitive brain research **23**(2-3) (2005) 436–443
14. Vicon Plug-in-gait: http://www.vicon.com/applications/gait_analysis.html
15. Demoucron, M., Askenfelt, A., Caussé, R.: Mesure de la "pression d'archet" des instruments à cordes frottées : Application à la synthèse sonore. In: Actes du VIIIème Congrès Français d'Acoustique, Tours. (2006)
16. Guettler, K.: Looking at starting transients and tone coloring of the bowed string. In: Proceedings of Frontiers of Research on Speech and Music. (2004)

ARTICLE IV

Effort-based analysis of bowing movements: evidence of anticipation effects

Nicolas Rasamimanana, Frédéric Bevilacqua
IRCAM, CNRS - UMR STMS,
1 Place Igor Stravinsky,
75004 PARIS, France
{Nicolas.Rasamimanana, Frederic.Bevilacqua}@ircam.fr

December 15, 2008

Abstract

Anticipatory behaviours are known to occur in music performance, notably on the control movements of instruments such as piano or drums. We studied such effects on bowed string movements, corresponding to a case where the control on sound is continuous. Movements were measured with an optical motion capture system combined with sensors on the bow. Bowing movements were analysed and compared on the basis of underlying effort costs, determined from their velocity profiles. Precisely, we used movement models that assume that jerk or impulse are minimized. These models were synthesized based on measurement data and then compared to velocity and acceleration profiles. Results on various musical cases involving separate strokes, scales, mixed bowing techniques and rhythms showed that this methodology can account, to some extent, for the different effort strategies used by the players. The presented modelling provides evidence of anticipatory behaviour during bowing movements.

1 Introduction

As evidenced by the extensive development of gesture-based interfaces for music and the increasing number of conference sessions dedicated to gesture controlled music, an important trend in computer music lies in the possibilities of controlling computer-generated sounds in real-time through the use of movements and physical gesture. This can enable electronic musicians to involve their body and motion in performance, recreating essential interactions found between musicians and their instruments (Leman, 2007). In that perspective, the study of different control strategies and constraints that acoustic instrument players must manage can bring important insights for such new approaches in electronic music performance.

A common assumption in music performance research is to consider music playing as a sequence of actions. Particular studies showed that in such motion sequences, each element can be affected by its neighbours. Pianists, for example, can modify their finger movements to anticipate the subsequent keystrokes one or two keystrokes ahead (Engel et al., 1997). Such an anticipatory movement behaviour can supposedly ease timing constraints inherent to music performance, e.g. tone production or rhythm, similarly to coarticulation phenomena found between phones in speech production (Engel et al., 1997; Godøy, 2004; Ortmann, 1929). From the analysis of drummers' playing, Dahl (2000) found that the performance of mixed stroke types influence stroke movements and timing. For example, in the case of patterns composed of one accented and three unaccented strokes, drummers typically prepare the accented stroke by raising their drumstick higher at the end of last unaccented stroke. They also lengthen the accent interval, which as a result gives more emphasis to the accent. These changes exemplify two underlying mechanisms that participate in anticipation, namely cognitive and biomechanical constraints. Analyses of piano player movements indeed revealed that cognitive chunking processes affect finger timings while biomechanical constraints have a notable effect on finger motion trajectories (Loehr and Palmer, 2007). Such anticipatory behaviours are in our view symptomatic of the different types of constraints (i.e. biomechanical, acoustical, cognitive, aesthetic) that acoustic instrument players must face to create expressive music.

The goal of this paper is to investigate anticipatory behaviours occurring with self-sustained instruments, like bowed strings or winds, where players have a continuous control on sound. On bowed strings for example, the production of sound results from the friction of bow hairs on the strings, moving at a certain velocity (Cremer, 1984). This acoustical mechanism enables bowed string players to continuously master their sound through a precise control of the bow. From this continuous control, they actually achieve different expressive cues as shown in bowing technique studies (Rasamimanana et al., 2006) or in analysing different performance versions (De Poli et al., 1998; Winold et al., 1994). In this perspective, it is interesting to study how possible anticipatory behaviours may happen in such instrument and investigate how anticipation challenges the instrument sound control movements. Previous studies showed some anticipation effects in violin playing, between left and right hands (Baader et al., 2005; Wiesendanger et al., 2006). While these studies examined coordination issues between fingering and bowing, we here aim to focus on anticipation occurring within bowing movements.

To carry out this study, we consider bowing movements as skilled movements, i.e. movements developed through training and practice to achieve certain objectives associated with a task. Under this assumption, bowing movements can be analysed under the scope of performance constraints optimization, and we can hence relate different effort costs to their execution.

Nelson (1983) introduced and described elementary principles underlying skilled movements. He defined different objectives related to physical economy of effort, e.g. minimizing energy cost or time cost, and showed that these objectives underlie different classes of skilled movements with specific velocity patterns. With this method, Nelson (1983) modeled and evaluated aspects of motor control strategies in jaw movements during speech. He also briefly reviewed violin bowing movements and pointed to possible effort costs. Links between underlying organizing principles and shapes of velocity patterns were also established by Hogan (1984) in the case of voluntary arm movements. Hogan (1984) proposed a mathematical model assuming smoothest possible movements to predict the velocity profiles of such movements. Inversely, it was also shown that the shape of velocity profiles can be related to effort costs as reported in (Ostry et al., 1987; Perkell et al., 2002) for speech movements.

We hence propose in this paper to compare different bowing movements on the basis of underlying effort costs. It was previously found that the execution of different bowing techniques, e.g. *Détaché* (sustained strokes) and *Martelé* (sharp almost percussive strokes), implies specific bow velocity profiles with characteristic bow accelerations at the start and end of strokes (Rasamimanana et al., 2006). Besides, different bow velocity patterns were found depending on tempo (Rasamimanana et al., 2007). These characteristics could interestingly be related to different efforts underlying bowing movements. To carry out this study, we analysed bowing movements using an analysis/synthesis method based on the kinematic formalism presented in (Nelson, 1983). First, this method allows us to study, quantify and interpret bowing variations due to adjacent strokes. Second, this approach also allows for the description of the bowing techniques reported in (Rasamimanana et al., 2006, 2007) with a different point of view.

This article starts with a brief presentation of the kinematic formalism and its adaptation to the study of bowed string movements. The measurements of real players' movements are then described. The analysis of measurements and the results are finally given and discussed.

2 Kinematic description of bowing movements

Instrument players achieve different *objectives*: on a first level, they aim at fulfilling objectives related to *music*, from technical aspects such as rhythm or tone production to expression. However, it was evidenced that on a second level, they also try to satisfy performance objectives that could be related to *physical economy* (Nelson, 1983). These musical and physical objectives define a set of constraints that influences the shapes of bow strokes velocity profiles. This section presents the kinematic formalism we used to describe bowing movements.

2.1 Formulation of physical constraints for bow control

We approximate bowing movement by a linear displacement of a mass m , along a dimension x . The mass spans the distance D in the movement time T , starts and ends with a null instantaneous velocity v . Moreover, an external force $F(t)$ is applied on mass m : the amplitude of this force is assumed to be bounded by a limit F_{max} . We consider $u(t) = F(t)/m$, generally referred as *control action* (Nelson, 1983), homogeneous to an acceleration and bounded by $U = F_{max}/m$. Considering a dissipative friction term $f_d(t)$, the system of equations describing the mass movement can be formalized as:

$$\begin{cases} \dot{x}(t) = v(t) \\ \dot{v}(t) = u(t) - f_d(t) \end{cases} \quad (1)$$

with the following conditions

$$\begin{cases} x(0) = 0, & x(T) = D, \\ v(0) = 0, & v(T) = 0, \\ |u(t)| \leq U \end{cases}$$

In this paper, we use this kinematic description for the longitudinal movements of the bow with respect to the instrument (violin, viola or cello). Adopting this formalism for each bow stroke, we obtain that strokes start and end with null velocity but not necessarily null acceleration (corresponding to bow changes), strokes have a given bow length D and duration T . The control action $u(t)$ actually corresponds to the acceleration that players give to the bow.

Moreover, one can notice that system (1) formulation is particularly straightforward. It should be stressed that similarly to (Nelson, 1983), the aim here is not to give a detailed and complete model of bowing movements, but rather to propose a simple and efficient description to evidence the relationships between physical dynamics, physical constraints, and performance objectives.

Furthermore, we assume an ideal movement without dissipative forces, i.e. $f_d(t) = 0$. This hypothesis is of course debatable in the case of movements related to bowed strings, since the vibration of the string is made possible from the friction force applied by the bow (Cremer, 1984). Nevertheless, on a first approximation, the friction characteristic can be modelled as a viscous friction (Serafin, 2004). This case is actually addressed by Nelson (1983, see appendices): although some changes actually appear on the analytical solutions of (1), especially on the absolute values of velocities, the principal characteristics of the velocity profiles remain unchanged. As a consequence, the relationships between physical aspects and performance objectives remain essentially the same. For this reason and for the sake of simplicity, dissipative forces are neglected as a first approach.

The model presented in system (1) is henceforth used to describe the kinematics of string bowing movements. We now introduce the concept of performance objectives and their consequence on the solutions of system (1).

2.2 General definition of performance objectives and impact on velocity profiles

To solve the equation system (1), a function has to span the distance D in the time T . There is actually an infinity of functions that may satisfy such conditions. However, as stated in (Nelson, 1983), solutions corresponding to skilled movements should also satisfy performance objectives that can be expressed as the minimization of a physical "cost" associated with the movement. This eventually defines supplementary global constraints. In this paper, we focus on the subclass of *unimodal* solutions, i.e. with velocity profiles whose slope sign changes exactly once: such movements indeed represent a certain efficiency of movement, since they correspond to a single accelerative phase and a single decelerative phase. Besides, such movements are also consistent with previously studied bow strokes (Rasamimanana et al., 2006). The next paragraph presents different performance objectives for such solutions.

Performance objectives are defined as minimizations of physical cost measures. They typically relate to *time*, *force*, *impulse*, *energy*, *jerk* (i.e. acceleration variations) and can be expressed analytically, as presented in appendix A. These costs actually define specific velocity profiles as shown on Figure 1 where different solutions are plotted for a fixed set of measured data: $D = 0.63m$, $T = 1s$ and $U = stroke_max_acceleration = 25m/s^2$. We can see that:

- Minimizing force (A), yields to a triangular pattern for instantaneous velocity, with slopes $(A_m, -A_m)$,
- The minimum impulse solution (I) has a trapezoidal shaped pattern with start and end slopes equal to U and $-U$ and minimum peak velocity V_m ,
- The minimizations of energy (E) and jerk (J) costs yield to velocity patterns with smooth bell shapes.

The analytical solutions satisfying (1) and the above performance objectives are given in appendix B.

Minimum impulse and minimum jerk are cost objectives particularly relevant to bowing movements considered in our study and are further detailed below.

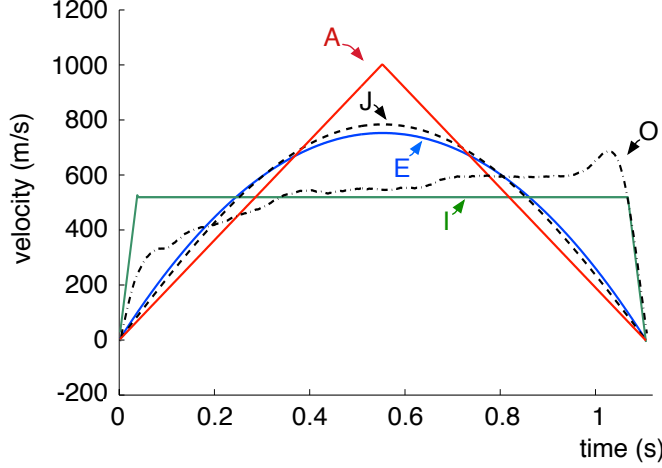


Figure 1: Velocity patterns minimizing different cost objectives. (O): original measured velocity pattern. (I): minimum impulse. (E): minimum energy. (J): discrete minimum jerk. (A): minimum force. Adapted from (Nelson, 1983)

2.3 Minimum impulse and minimum jerk objectives

Minimum impulse solutions minimize the total impulse, i.e. the time integral of the control action function u , over the stroke, given D , T , and U . However, we can also consider these solutions as minimizing the velocity variations over the stroke given D , T and U . Solving system (1) under this constraint indeed yields a trapezoidal velocity profile where the velocity variations are $+U$, 0 , and $-U$: this form has often been reported as "*Bang-Zero-Bang*". In this perspective, these solutions can be particularly appropriate to describe bowings: players can keep a relatively constant bow velocity to obtain a sustained sound.

Minimum jerk movements, which minimize acceleration transients, are widely used in motor control studies to describe free human movements. Two types of minimum jerk solutions can be defined according to the considered movements, i.e. discrete or cyclical (Nelson, 1983; Hogan and Sternad, 2007). Discrete movements are characterized with well defined start and stop phase, and are separated by a pause. Thus, discrete minimum jerk solutions specifically impose null velocity and null acceleration at the beginning and end of each movement. Cyclical movements correspond to sequences of repeating patterns with no pause. In this case, cyclical minimum jerk solutions have null velocity and non-null acceleration at the beginning and end of each movement. Moreover, as we consider repeated "back and forth" movements, acceleration is maximum at those moments. These two

types of solutions have distinct velocity patterns, both characterized by a bell shape. These solutions are relevant for bowing in the case of stopped or repeated strokes.

Using measured times T , bow lengths D , bow accelerations u from real players data and the analytical solutions to system (1), different bow velocity patterns can be synthesized satisfying different performance objectives. It is then possible to assess performance objectives involved in the playing by comparing synthesized and measured velocity profiles.

3 Setup and procedure

This section describes the setup, protocole and musical material used for the study.

3.1 Sound and movement measurements

We used a Vicon System 460 optical motion capture system to measure bow motion. Six M2 cameras were placed around the instrumentalist, providing a spatial resolution below 1mm at a frame rate of 500Hz, on a volume of approximately $1m^3$. Six markers were placed on the instrument, four on the violin table, one on the nutmeg and one on the tailpiece to indicate the strings position. Three markers were placed on the bow. Figure 2 shows the marker placement.

An additional 3 axis ADXL202 accelerometer was fixed at the frog of the bow. Accelerometer data was digitized at 500Hz and transmitted to a laptop for recording. To guarantee post-recording synchronization between motion capture data and accelerometer data, the sound track was recorded simultaneously by each sensing system. The coupling of motion capture data and accelerometer data grants a precise time and space measurement of players position, velocity and acceleration (Rasamimanana, 2008).

The markers and sensors placed on the bow added less than two grams, mainly at the frog (the bow was 62g and the frog was 17g): the players did not feel being disturbed by the system.

3.2 Procedure

We asked three bowed string players to perform five different musical situations. All were advanced level players with eight to ten years of practice. Five tasks were asked:

- *Task1*: the violinists played series of ten isolated strokes, i.e. quarter notes interleaved with pauses, in *Détaché* and in *Martelé* (see Figure 5 for music score).

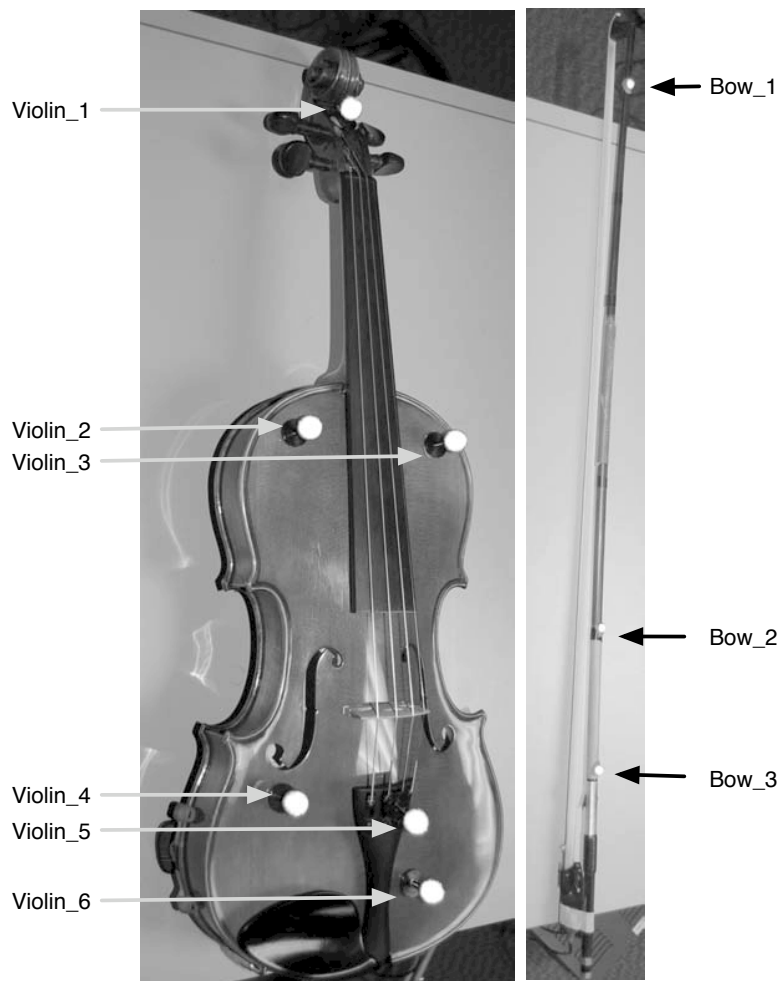


Figure 2: Marker placement to measure violin and bow movements through motion capture.

- *Task2*: the violinists played one ascending and descending scale (one octave), in *Détaché* and in *Martelé* (see Figure 5 for music score).
- *Task3*: the violinists played an exercise mixing *Détaché* and *Martelé* bowings. It consisted in a pattern of four quarter notes where the first two notes are *Détaché* and the last two notes are *Martelé*. The reverse pattern, first two notes *Martelé* and last two note *Détaché*, was also recorded. Each pattern was repeated six times, therefore constituting a set of 36 patterns mixing the two bowings (see Figure 6 for music score).
- *Task4*: the violinists played an exercise mixing different rhythms. The rhythmic pattern was constituted of a quarter and four sixteenth notes, repeated four times on different notes (see Figure 7 for music score).
- *Task5*: the violinists performed a *Détaché accelerando* from moderate tempo (80 bpm) to "as fast as possible", tied with a *decelerando* (see Figure 8 for music score). This exercise was the same as reported in (Rasamimanana et al., 2007).

No specific indications were given for bowing directions: violinists alternated downbows and upbows. For the tasks 1 to 4, the players were asked to play at a moderate tempo, 80 bpm and at a *forte* dynamic.

It should further be noticed that we did not aim here to perform statistics on a large number of players, as a great variability would be expected (Winold et al., 1994; Dahl, 2000; Rasamimanana, 2008). We rather focused on an in depth study of a smaller number of expert instrumentalists, taking into account each idiosyncratic playing.

4 Method

This section details the analysis/synthesis process we develop in this study. First we present the synthesis method for cost-based velocity profiles based on measured data features. Second, we present the method of computing differences between measured and synthesized profiles, which permits quantitative assessment of underlying effort costs.

4.1 Computation of bowing parameters

The first step corresponds to extracting bowing parameters from measured data. The markers on the violin are used to define a frame of reference. The transverse displacement of the bow, $x(t)$, is then calculated in this frame from the projection of bow marker positions. Motion derivatives of the bow displacement, $v(t) = \frac{dx(t)}{dt}$ and $a(t) = \frac{d^2x(t)}{dt^2}$ are then obtained from differentiation and smoothing, using a Savistky-Golay filter. The parameters

of the filter are manually adjusted by comparing the acceleration obtained from the motion capture to the signal from the accelerometers: for moderate tempo exercises, the order was set to 3 and the window size to 80ms.

As defined previously, each bow stroke is segmented based on the velocity zero crossings: a stroke starts and ends with null velocity. For each stroke, we compute:

- the used bow length $D = |x(t_{k+1}) - x(t_k)|$, where t_k are the instants of zero velocity,
- the movement duration $T = t_{k+1} - t_k$,
- the initial bow acceleration $a_0 = a(t_k)$,
- the maximum and minimum bow acceleration $a_{max} = \max_{t \in [t_k; t_{k+1}]} a(t)$,
 $a_{min} = \min_{t \in [t_k; t_{k+1}]} a(t)$.

These computed data are then used as empirical conditions to synthesize bow velocity profiles.

4.2 Synthesis of bow velocity profiles

For each stroke, three velocity patterns satisfying system (1) are synthesized using the computed bowing parameters. These models are:

- a trapezoidal model Tr , derived from minimum impulse solutions,
- a discrete minimum jerk model J_d ,
- a cyclical minimum jerk model J_c .

The Tr pattern is a deformable trapezoidal shape. The initial and terminal slopes are determined by a_{min} and a_{max} and the constant peak velocity is computed taking into account bow length D , time T and the slopes. The full analytical expression of the model is derived from minimum impulse solutions and can be found in appendix C. Like minimum impulse solutions, this model minimizes the total impulse (time integral of the control action) over the stroke. However, the minimization is here computed based on conditions determined by the measured initial and terminal accelerations, instead of a constant maximum acceleration value. For this trapezoidal model, the velocity variations are described by the triplet $a_{max}, 0, a_{min}$.

The discrete and cyclical minimum jerk solutions, J_d and J_c respectively, minimize acceleration transients over the stroke (see section 2.3). These two solutions can be expressed as polynomials¹ and full analytical expressions

¹The analytical expressions for minimum jerk models are generally found from dynamical optimization.

are given in appendix C. The discrete model J_d and cyclical model J_c are computed taking into account the initial acceleration a_0 , the maximum acceleration over the stroke a_{max} , the distance spanned D and the duration of movement T . In this process, an additional condition must be defined to comply with the unimodal constraint we chose: acceleration curves can not change sign more than once per stroke. We consider that the minimization has no solution when only non-unimodal solutions are possible with the measured data.

It must be noticed that when a_0 approaches the maximum acceleration over the stroke a_{max} , the velocity model of discrete minimum jerk J_d approaches the cyclical minimum jerk model J_c . This model can therefore account for continuity between discrete and cyclical movements (Nelson, 1983; Hogan and Sternad, 2007).

4.3 Comparison function

At this point, a comparison function determining which synthesized model is the most similar to the measured data is needed. In (Nelson, 1983; Ostry et al., 1987; Perkell et al., 2002), the velocity patterns are compared on the basis of their maximum peak velocity normalised by the average velocity. However, this single value may not be sufficient to determine convincingly similarity between velocity patterns. For example, on Figure 1, all velocity patterns have identical average velocity. The measured data (O) looks trapezoidal and therefore, based on its shape, the most similar model should be minimum impulse (I). However, looking at peak velocities, it appears that (O) should be associated with minimum energy (E), and not minimum impulse (I).

In this paper, we propose a different approach. We base our analysis on the entire profile by considering correlations between data and model curves, over each strokes. Correlation here accounts for linear relationships between the two curves and therefore stresses velocity pattern shape similarities. In this case, the measured data (O) is eventually associated with the minimum impulse model (I) as they are the most correlated. Moreover, since it was found in (Rasamimanana et al., 2006) that acceleration curves show characteristic properties depending on bowing techniques, we choose to carry out our analysis on acceleration profiles.

Figures 3 and 4 show the synthesized velocity patterns computed from the trapezoidal model and the minimum jerk model, their corresponding accelerations curves, and the measured data for strokes in *Détaché* and in *Martelé*.

On the graphs, one can observe that even when good match can be found between the data and the model, e.g. on Figure 4 between the minimum jerk model and data, there remains some differences. This can be explained by the fact that the cost based minimization represents only one of the different

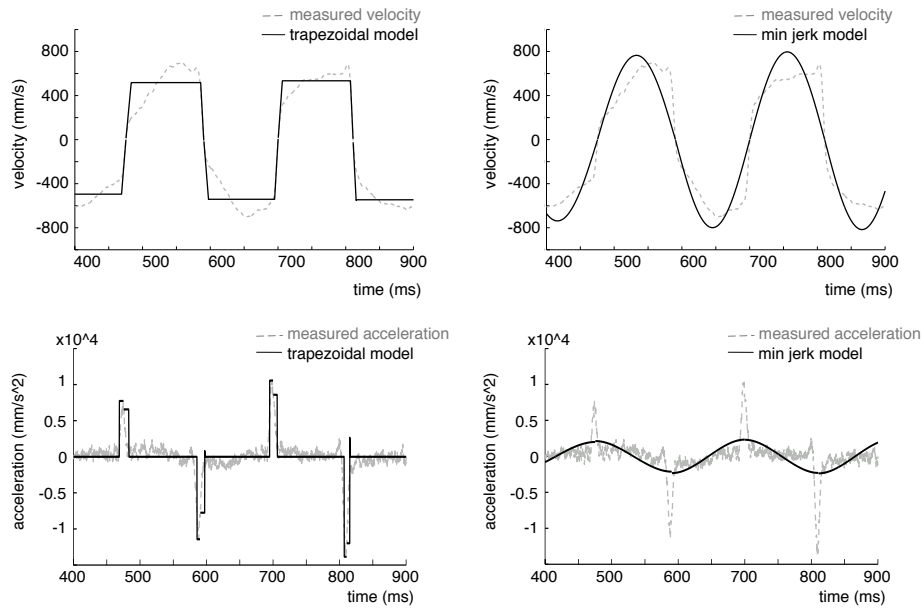


Figure 3: Trapezoidal and cyclical minimum jerk models for four measured *Détaché* strokes. Top plots show the measured and the synthesized velocity patterns. Bottom plots show the acceleration patterns. Positive and negative velocities show alternation between downbows and upbows

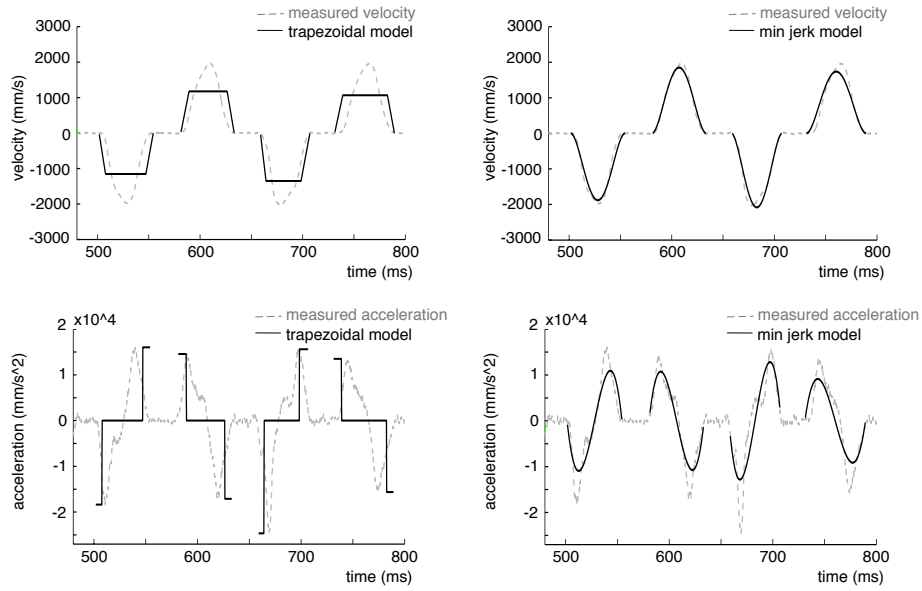


Figure 4: Trapezoidal and cyclical minimum jerk models for four measured *Martelé* strokes. Top plots show the measured and the synthesized velocity patterns. Bottom plots show the acceleration patterns. Positive and negative velocities show alternation between downbows and upbows

constraints that players must deal with when bowing. Another reason that could be pointed out is the assumption of unimodal velocity profiles, which imposes possible velocity variations to a single increasing phase and a single decreasing phase.

The comparison process is performed as follows. For each bow stroke, we correlate the acceleration curve to each of its synthesized versions (Tr , Jd and Jd). This hence gives a triplet of correlation factors that accounts for the similarity between a single measured stroke and its associated models. The model giving the highest correlation value corresponds to the most similar model. Besides, it should be noticed that by providing separate measures for all three models, this triplet actually reflects possible trade-offs between performance objectives, as introduced in (Nelson, 1983).

5 Results

This section begins with the analysis of velocity profile shapes in simple cases, i.e. isolated notes and scales, for the bowing techniques *Détaché* and *Martelé*. The results are then compared to more complex cases, such as mix of fast and slow rhythms, *accelerando* / *decelerando* and an exercise with mixed bowing techniques. The evolution of velocity profile shapes according to the bow stroke context is evaluated using the previously described method.

5.1 Single strokes and scales

In a first step, bow strokes are studied in simple, stereotypical cases involving the two bowing techniques *Détaché* and *Martelé*.

5.1.1 Single strokes

The first case corresponds to the performance of individual strokes separated with silence at a moderate tempo, namely quarter notes at 80 bpm, at a *forte* dynamic, interleaved with pauses (*Task1*). For each stroke, the three previously described velocity models are computed and compared to the measured data using the acceleration curves correlation. The results show that for all players, the most similar model to the measured separate strokes is the discrete minimum jerk model J_d , for both bowing technique *Détaché* and *Martelé*. These movements therefore appear to satisfy an optimal smoothness over the execution of strokes for both bowing techniques. This result actually coincides with classical results on voluntary arm movements in one dimension, generally minimizing jerk over the task, such as described for example in (Hogan, 1984) in the case of reaching movements.

The comparison results are illustrated for one player on Figure 5 for *Détaché* strokes (top left) and for *Martelé* strokes (top right). On top, the

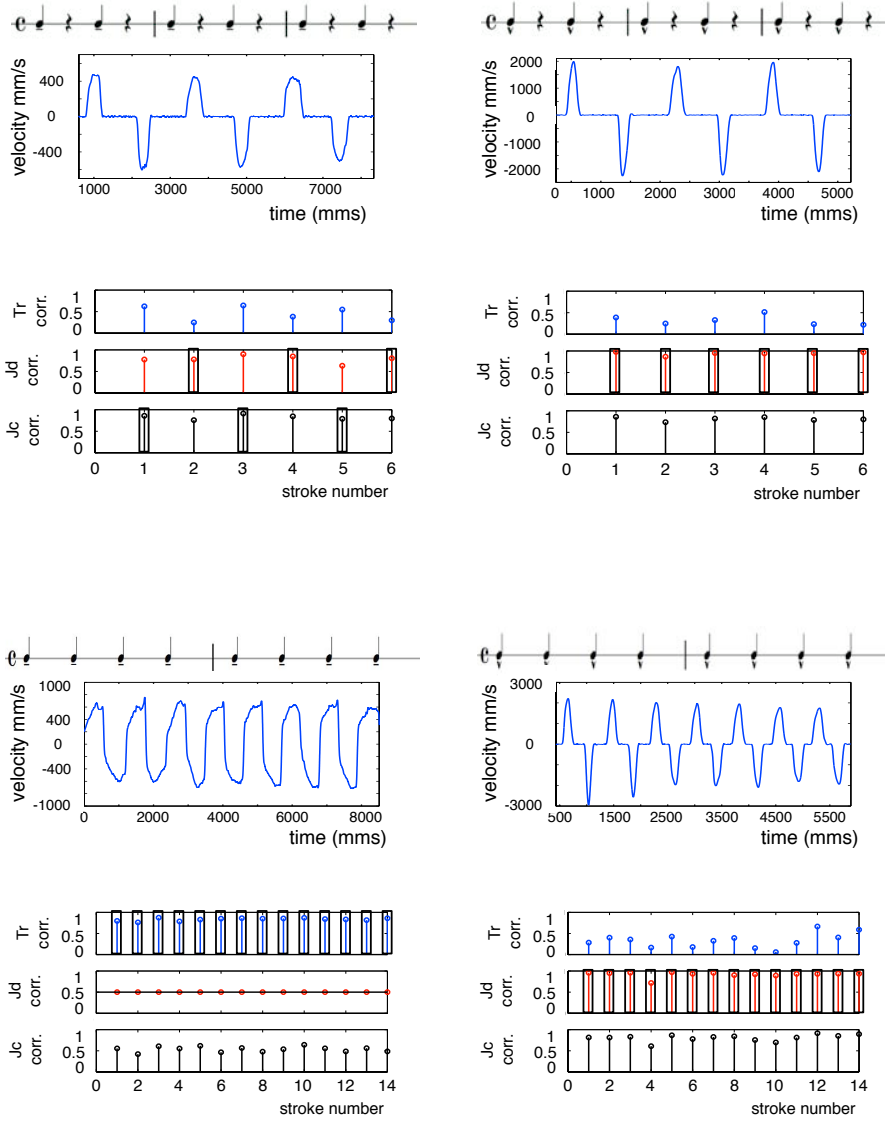


Figure 5: for all plots, from top to bottom: score, measured velocity patterns and correlation factors for the three models (Tr , Jc and Jd). For each stroke, the highest correlation factor is highlighted with a rectangle. Plots on the left correspond to *Détaché* and plots on the right correspond to *Martelé*.

measured velocity curve is displayed. The graphs below shows the plot of the correlation factors for the trapezoidal model Tr , the discrete minimum jerk J_d and the cyclical minimum jerk J_c according to each stroke. For each stroke, the highest correlation factor is highlighted (the stroke is surrounded by a rectangle). For both bowing techniques, the correlation factors of the trapezoidal model are much lower (< 0.6) than the correlation factors for discrete and cyclical minimum jerk (> 0.85). In the case of *Détaché*, the discrete and cyclical jerk models are particularly similar, as expressed by the correlation factors, showing that the initial accelerations are close to the maximum accelerations.

5.1.2 Scales

The second stereotypical situation corresponds to scales performed in *Détaché* and in *Martelé* at a moderate tempo (80 bpm) at a *forte* dynamic (*Task2*). This case allows for the study of possible differences between the performance of discrete, separate strokes and a continuous series of bow strokes. The results actually show interesting differences from single strokes.

For *Détaché*, the best model that describes the measured velocity is the trapezoidal model Tr as displayed on Figure 5 bottom left. This is different from the single strokes case where J_d represented the best model. For all three bowed string players, the correlation factors are indeed the highest for this model (with very high values between 0.85 and 0.9) compared to the others. It is also interesting to note that in this case, the discrete minimum jerk model J_d can not yield to any possible unimodal solutions for the strokes. The cyclical minimum jerk model J_c provides a less appropriate fit with the measured data with correlation factors around 0.5. In this case, the movements involved in the performance of *Détaché* are not optimally smooth anymore, but appear to be better explained with minimum velocity variations over the stroke.

To the contrary, for *Martelé*, the results are similar to the discrete case, as the model that best correlates with the data still is the discrete minimum jerk pattern J_d . However, as can be seen for *player1* on Figure 5 bottom right, players generally stopped the bow between strokes when performing a *Martelé* scales. This short pause is indeed used to put some pressure on the bow before the beginning of strokes, a characteristic of this bowing technique (Demoucron et al., 2006): formally, these *Martelé* scales corresponds to a sequence of discrete movements.

The analyses of bow velocity profile shapes for these simple bow strokes are used as references to study the more complex situations, as described in the following sections.

5.2 Mixed bowing techniques

We consider in this section a mix of two types of bowing techniques, similarly to the study by Dahl (2000) on drums. The musical situation consists in a fragment of a violin study made of patterns of four quarter notes, with two *Détaché* and two *Martelé*, performed at a moderate tempo (80bpm) and at a *forte* dynamic (*Task3*). In the previous section, we found that, in the case of scales, *Détaché* is best described with the trapezoidal model, while *Martelé* is best described with the discrete minimum jerk model. Therefore, based on this result, the sequence of four quarter notes should lead to (Tr , Tr , J_d , J_d). The performed analysis however shows a different result as shown on Figure 6. For each stroke, statistics are reported over the 36 recorded patterns performed by 3 players. Figure 6 shows the percentage of each model for the four strokes of a pattern. The first, third and fourth strokes mainly corresponds to the expected models, i.e. Tr (76.5%), J_d (76.5%) and J_d (100%), respectively. Interestingly, the second stroke is best

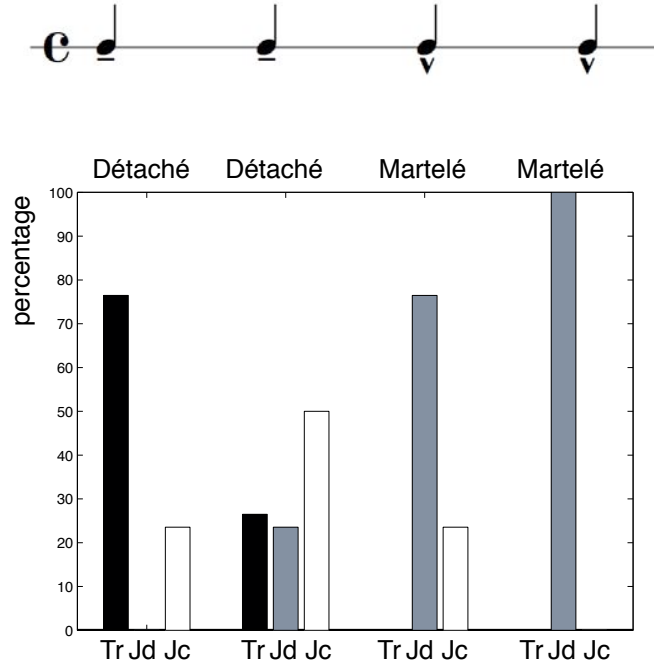


Figure 6: Mixed bowing techniques: percentage of each model for the four strokes of a pattern, consisting of two *Détaché* and two *Martelé*. Results include cases reversing *Martelé* and *Détaché*.

explained by the cyclical minimum jerk movement J_c (50%), instead of the expected trapezoidal movement Tr (26.5%). In other words, when consid-

ering a sequence of four *Détaché*, like in the case of scales, all bow strokes are best explained with movements minimizing the velocity variations (*Tr*). However, when mixing two *Détaché* and two *Martelé* in a sequence of four strokes, the second stroke appears to be affected by the models that was found to corresponds best to *Martelé*, i.e. optimally smooth movements. This can be viewed as a property of anticipatory behaviours.

Inversely, when considering a sequence of four *Martelé*, e.g. in the case of scales, the bow strokes are all best explained with discrete minimum jerk movements J_d , the bow generally being stopped before a *Martelé*. In this sequence of four strokes, however, it can be noticed that the third stroke, i.e. the first *Martelé*, is in some cases best explained by J_c (23.5%). This therefore indicates that in those cases, the bow is not completely stopped between the end of the *Détaché* stroke and the beginning of the *Martelé* stroke. This result could actually be interpreted as a phenomenon of persistence of the last *Détaché* stroke on the first *Martelé* stroke.

In speech, anticipation and persistence between phones are the two aspects that define *coarticulation*. It can be particularly interesting here to draw such a parallel between speech production and bowing movements. More specifically, coarticulation in speech is known to be essential for intelligibility, although not always directly audible. We can wonder whether this *gesture coarticulation* could contribute to a *musical intelligibility*, giving to a performance some organic properties that makes it humanly plausible, as opposed to purely computational processes.

5.3 Mixed rhythms and influence of stroke frequency

The following musical exercises extend the previous cases by the addition of different rhythms.

5.3.1 Mixed rhythms

The performers played a series a rhythmic patterns, consisting in a quarter note followed by four sixteenth notes, at a moderate tempo (80bpm), at a *forte* dynamic level (*Task4*). This situation therefore creates a context mixing slow and fast rhythms. The same analysis is performed. Results are close for player1 and player2 as discussed below. Player3 is discussed separately.

Results for player1 and player2 indicate that the effort-based models actually depend on the rhythms. All quarter notes display a similar bow velocity profile that is best described with the trapezoidal model *Tr*, while sixteenth notes all have velocity profiles best described with the cyclical minimum jerk model J_c . Namely, the performance of this rhythmic pattern is best described with the sequence (*Tr*, J_c , J_c , J_c , J_c , *Tr*). This result actually suggests that different performance objectives underlie the execu-

tion of this sequence of rhythms. On quarter notes, players 1 and 2 tend to minimize the variation of velocity, producing a relatively constant sound, as opposed to sixteenth notes, where players 1 and 2 tend to produce the smoothest movement for the four of them. This result therefore invites to distinguish two *musical gestures* in their performance, i.e. one for the quarter notes and one for the sixteenth notes, which is actually found consistent from the players' viewpoint: intuitively, four sixteenths are not played as four times one sixteenth but rather as one unit. Figure 7 displays the correlation results for a performance of four rhythmic patterns. The results show patterns of one trapezoidal model and four minimum jerk models.

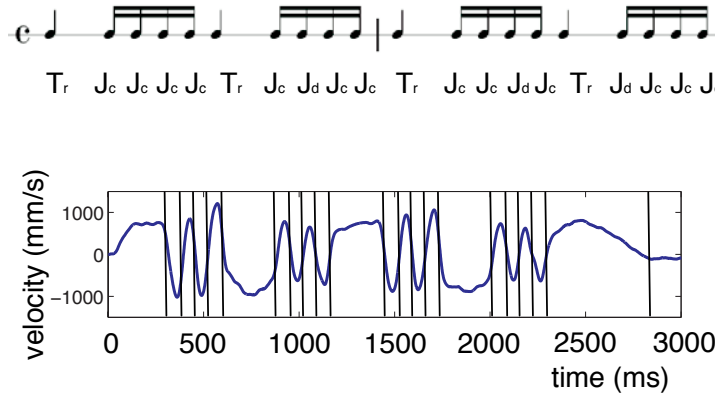


Figure 7: Mixed rhythm: score and measured velocity patterns. The best explaining model is given for each stroke.

Interestingly, the results for player3 are relatively different. The strokes are all closer to the cyclical minimum jerk model. In the light of our previous results, this could suggest that the player performed the sequences of rhythmic patterns as one whole. These results on the three players clearly show that in a musical context, the instrumentalists have different proper performance strategies. Such differences were not found in the case of single strokes or scales.

5.3.2 Accelerando / Decelerando

The next bow stroke context considers an *accelerando* / *decelerando*, i.e. a series of *Détaché* strokes performed with an increasing frequency followed by a series of strokes with a decreasing frequency (*Task5*). This case was already studied in the article (Rasamimanana et al., 2007), where different velocity patterns were reported and characterized with a sinusoidal non-linear fit. The analysis done in this article enables to further understand

these results as displayed on Figure 8. As in the case of the above mixed rhythm sequence, slowest strokes are best described by the trapezoidal model and fastest strokes by minimum jerk models, therefore approving that the fastest movements tend to optimize acceleration smoothness and that the slowest movements tend to optimize velocity variation. Interestingly, the frequencies of the changes from one cost to the other are similar to those reported in (Rasamimanana et al., 2007), i.e. around 8Hz.

Globally, this modelling confirms the existence of the transitions reported in (Rasamimanana et al., 2007). Moreover, the modelling presented here brings a new interpretation: the slow and fast parts of the *accelerando* / *decelerando* can be explained by different optimizations, from minimization of velocity variations to minimization of jerk.

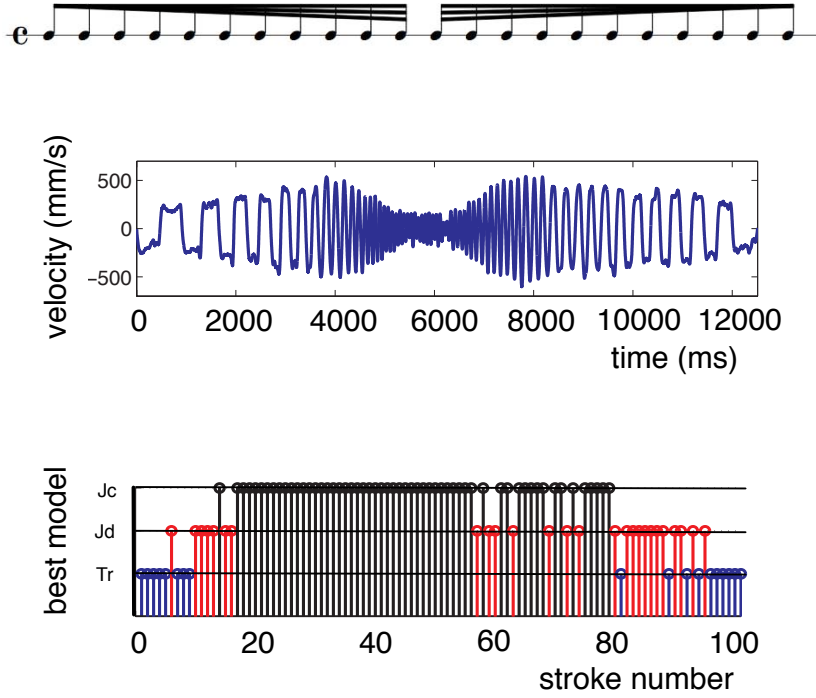


Figure 8: Accelerando / Decelerando: score, measured velocity patterns and best models (highest correlation).

6 Summary and perspectives

In this article, we considered effort costs associated with bowing movements to analyse possible strategies expert players use. With the help of a kinematic formalism, we related bowing techniques to different effort costs and based on this methodology, we evidenced the possible influence of the bow stroke context, in particular the influence between strokes.

The results on single strokes and scales show that the different bowing techniques can be modelled as different performance optimizations. Moreover, different optimizations also depend on tempo as shown with the *accelerando* / *decelerando*. At least two minimization costs are therefore relevant, namely jerk minimization and minimization of velocity variations. Nevertheless, notable differences between players are found depending on the musical context: mixing rhythms or bowing techniques. Such differences are actually expected since musicians might use different bowing strategies to play a given musical passage (Winold et al., 1994). The methodology we proposed in this paper seems to be able to take into account such variations, providing a framework for the interpretation of the underlying strategies. For example, we were able to evidence from a gesture perspective different groupings of notes related to phrasing. Moreover, the modelling presented in this article enabled to show anticipatory behaviour and coarticulation on bowing movements: the bowing movement of strokes is influenced by its neighbors, as occurring in speech coarticulation between phonemes. We believe that the evidence of such behaviour is important since it reveals a promising path to improve playability of digital musical instruments that usually do not take into account such effects.

More generally, the consideration of effort constraints can bring interesting insights to music performance analysis and synthesis. This approach can for example be useful for the definition of sound controls, still largely based on a MIDI/ADSR framework. The results presented in this paper indeed indicate that designing control as successive independent units is limiting compared to music playing on acoustic instruments. Moreover, this approach can also be useful to simplify the control of sound synthesis. Classes of temporal profiles related to physical costs can be intuitively controlled with only few input parameters.

7 Acknowledgements

We would like to acknowledge Pr. Paolo Viviani for interesting discussions. We also would like to thank Marcelo Wanderley, Matthias Demoucron, Norbert Schnell and René Caussé for contributions to this work and for fruitful collaborations. We also acknowledge support of the project CONSONNES funded by the ANR (Agence Nationale de la Recherche, France). This work

was partly funded by the european project i-Maestro IST-026883 (www.i-maestro.org).

A Measures of physical cost

Five different performance costs.

$$\text{time cost: } T = \text{movement time}, \quad (2)$$

$$\text{force cost: } A = \max_{t \in (0, T)} |u(t)|, \quad (3)$$

$$\text{impulse cost: } I = \frac{1}{2} \int_0^T |u(t)| dt, \quad (4)$$

$$\text{energy cost: } E = \frac{1}{2U} \int_0^T u^2(t) dt, \quad (5)$$

$$\text{jerk cost: } J = \int_0^T \dot{a}^2(t) dt, \quad (6)$$

where, $u(t)$ is the control action, U is the control action limit, $a(t)$ is the mass acceleration, and T the duration of movement.

B Analytical solutions

In the case of negligible frictions, the analytical solutions to system 1 are:

$$\begin{aligned} \text{minimum time:} \quad V_t(t) &= \begin{cases} Ut & \text{if } t \in [0; \frac{T_m}{2}[\\ -Ut & \text{if } t \in [\frac{T_m}{2}; T_m] \end{cases} \quad \text{where,} \\ T_m &= \sqrt{4D/U}, \end{aligned}$$

$$\begin{aligned} \text{minimum force:} \quad V_f(t) &= \begin{cases} A_m t & \text{if } t \in [0; \frac{T}{2}[\\ -A_m t & \text{if } t \in [\frac{T}{2}; T] \end{cases} \quad \text{where,} \\ A_m &= 4D/T^2, \end{aligned}$$

$$\begin{aligned} \text{minimum impulse:} \quad V_m(t) &= \begin{cases} Ut & \text{if } t \in [0; \frac{V_0}{U}[\\ V_0 & \text{if } t \in [\frac{V_0}{U}; T - \frac{V_0}{U}[\\ V_0 - Ut & \text{if } t \in [T - \frac{V_0}{U}; T] \end{cases} \quad \text{where,} \\ V_0 &= \frac{TU}{2} - \sqrt{(\frac{TU}{2})^2 - DU}, \end{aligned}$$

$$\text{minimum energy:} \quad V_e(t) = 6D/T^2 (t - t^2/T),$$

$$\begin{aligned} \text{discrete minimum jerk:} \quad V_{jd}(t) &= A_0 t \left[1 + 6 \left(\frac{5D}{A_0 T} - 1 \right) \frac{t}{T} - 5 \left(\frac{6D}{A_0 T^2} - 1 \right) \left(2 \frac{t^2}{T^2} - \frac{t^3}{T^3} \right) \right], \\ &\text{where } A_0 \text{ is the initial acceleration.} \end{aligned}$$

C Velocity models

trapezoidal model:
$$V_{trap}(t) = \begin{cases} a_{max} t & \text{if } t \in [0; \frac{V_p}{a_{max}}[\\ V_p & \text{if } t \in [\frac{V_p}{a_{max}}; T - \frac{V_p}{|a_{min}|}] \\ V_p - |a_{min}|t & \text{if } t \in [T - \frac{V_p}{|a_{min}|}; T] \end{cases}$$

where, $V_p = \left(\frac{1}{a_{max}} + \frac{1}{a_{min}} \right) \left(T - \sqrt{T^2 - 2D \left(\frac{1}{a_{max}} + \frac{1}{a_{min}} \right)} \right)$,

discrete minimum jerk:
$$V_{jd}(t) = A_0 t \left[1 + 6 \left(\frac{5D}{A_0 T} - 1 \right) \frac{t}{T} - 5 \left(\frac{6D}{A_0 T^2} - 1 \right) \left(2 \frac{t^2}{T^2} - \frac{t^3}{T^3} \right) \right],$$

where A_0 is the initial acceleration.

cyclical minimum jerk:
$$V_{jc}(t) = 5D/T^2 \left(t - 2 \frac{t^3}{T^2} - \frac{t^4}{T^3} \right),$$

References

- Baader, A. P., Kazennikov, O., and Wiesendanger, M. (2005). Coordination of bowing and fingering in violin playing. *Cognitive brain research*, 23(2-3):436–443.
- Cremer, L. (1984). *The Physics of the Violin*. The MIT Press.
- Dahl, S. (2000). The playing of an accent - preliminary observations from temporal and kinematic analysis of percussionists. *Journal of New Music Research*, 29:225–233.
- De Poli, G., Rodà, A., and Vidolin, A. (1998). Note-by-note analysis of the influence of expressive intentions and musical structure in violin performance. *The Journal of New Music Research*, 27(3):293–321.
- Demoucron, M., Askenfelt, A., and Caussé, R. (2006). Mesure de la "pression d'archet" des instruments à cordes frottées : Application à la synthèse sonore. In *Actes du VIIIème Congrès Français d'Acoustique, Tours*.
- Engel, K., Flanders, M., and Soechting, J. (1997). Anticipatory and sequential motor control in piano playing. *Experimental Brain Research*, 113:189–199.
- Godøy, R. I. (2004). Gestural imagery in the service of musical imagery. In Camurri, A. and Volpe, G., editors, *Lecture Notes in Artificial Intelligence, LNAI 2915*, pages 55–62. Springer Verlag.
- Hogan, N. (1984). An organizing principle for a class of voluntary movements. *The Journal of Neuroscience*, 4(11):2745–2754.
- Hogan, N. and Sternad, D. (2007). On rhythmic and discrete movements: reflections, definitions and implications for motor control. *Experimental Brain Research*, 181:13–30.

- Leman, M. (2007). *Embodied music cognition and mediation technology*. The MIT Press.
- Loehr, J. D. and Palmer, C. (2007). Cognitive and biomechanical influences in pianists' finger tapping. *Experimental Brain Research*, 178:518–528.
- Nelson, W. L. (1983). Physical principles for economies of skilled movements. *Journal Biological Cybernetics*, 46(2):135–147.
- Ortmann, O. (1929). *The physiological mechanics of piano technique*. Dutton, New York.
- Ostry, D., Cooke, J., and Munhall, K. (1987). Velocity curves of human arm and speech movements. *Experimental Brain Research*, 68:37–46.
- Perkell, J., Zandipour, M., Matthies, M., and Lane, H. (2002). Economy of effort in different speaking conditions. i. a preliminary study of intersubject differences and modeling issues. *The Journal of the Acoustical Society of America*, 112(4):1627–1641.
- Rasamimanana, N. H. (2008). *Geste instrumentale du violoniste en situation de jeu : analyse et modélisation*. PhD thesis, Université Paris 6 - IRCAM UMR STMS.
- Rasamimanana, N. H., Bernardin, D., Wanderley, M., and Bevilacqua, F. (2007). String bowing gestures at varying bow stroke frequencies: A case study. In *Advances in Gesture-Based Human-Computer Interaction and Simulation*, volume 5085 of *Lecture Notes in Computer Science*, pages 216–226. Springer Verlag.
- Rasamimanana, N. H., Fléty, E., and Bevilacqua, F. (2006). Gesture analysis of violin bow strokes. In *Gesture in Human-Computer Interaction and Simulation*, volume 3881 of *Lecture Notes in Computer Science / Artificial Intelligence (LNAI)*, pages 145–155. Springer Verlag.
- Serafin, S. (2004). *The sound of friction: real-time models, playability and musical applications*. PhD thesis, Stanford University, CCRMA.
- Wiesendanger, M., Baader, A. P., and Kazennikov, O. (2006). Fingering and bowing in violinists: a motor control approach. In Altenmüller, E., Wiesendanger, M., and Kesselring, J., editors, *Music, Motor Control and the Brain*, pages 109 – 123. Oxford University Press.
- Winold, H., Thelen, E., and Ulrich, B. D. (1994). Coordination and control in the bow arm movements of highly skilled cellists. *Ecological Psychology*, 6(1):1–31.

Troisième partie

Annexes

ARTICLE NIME 2006A

COMBINING ACCELEROMETER AND VIDEO CAMERA :
RECONSTRUCTION OF BOW VELOCITY PROFILES.
E. SCHOONDERWALDT, N. RASAMIMANANA AND F. BEVILACQUA

Combining accelerometer and video camera: Reconstruction of bow velocity profiles

Erwin Schoonderwaldt
CSC
KTH
Lindstedtsv. 24
S-100 44 Stockholm

(+46) 8 790 7561
schoondw@speech.kth.se

Nicolas Rasamimanana
IRCAM
CNRS STMS UMR 9912
Real Time Applications Team
1, Place Igor Stravinsky
75004 Paris

(+33) 1 44 78 48 26
Nicolas.Rasamimanana@ircam.fr

Frédéric Bevilacqua
IRCAM
CNRS STMS UMR 9912
Real Time Applications Team
1, Place Igor Stravinsky
75004 Paris

(+33) 1 44 78 48 31
Frederic.Bevilacqua@ircam.fr

ABSTRACT

A cost-effective method was developed for the estimation of the bow velocity in violin playing, using an accelerometer on the bow in combination with point tracking using a standard video camera. The video data are used to detect the moments of bow direction changes. This information is used for piece-wise integration of the accelerometer signal, resulting in a drift-free reconstructed velocity signal with a high temporal resolution. The method was evaluated using a 3D motion capturing system, providing a reliable reference of the actual bow velocity. The method showed good results when the accelerometer and video stream are synchronized. Additional latency and jitter of the camera stream can importantly decrease the performance of the method, depending on the bow stroke type.

Keywords

Bowing gestures, bowed string, violin, bow velocity, accelerometer, video tracking.

1. INTRODUCTION

Accelerometers and standard video camera are two different types of widely used sensors in the design of cost-effective gesture capture systems. In particular, such sensors have been incorporated in several musical interfaces. Each of these types of sensor has different characteristics. First, accelerometers are typically used to build miniature low latency systems. They are for example particularly well suited to capture percussive gestures. Nevertheless, quantitative use of accelerometer might be difficult due to the fact that the signal depends on both the tilt angle and the actual acceleration. Second, video cameras are well suited to localize, and to spatially follow object. Nevertheless standard video rate are relatively slow for musical application and important latencies are difficult to avoid.

This very brief description points out that accelerometers and video camera are actually complementary systems [8]. Moreover, as discussed by Foxlin [7], there has been a growing interest in the field of Augmented Reality to combine both inertial systems (accelerometers and/or gyroscopes) and vision systems to perform efficient tracking. Generally, the inertial component is fixed to the camera.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

NIME 06, June 4-8, 2006, Paris, France.
Copyright remains with the author(s).

We report here a simple approach combining accelerometers with a standard camera to capture bowing gestures. Such a combination is interesting in this case since bowing gesture contains fast and slow temporal features, as well as small and large spatial displacements. As a matter of fact, in bowed string instruments such as the violin, bowing gestures form an essential part in the tone production, giving the player a continuous, yet complex, control of the produced sound.

The capture of bowing gestures have shown important interests in various contexts. Bowing measuring systems can be either used in *non realtime* for fundamental studies of violin playing or as gestural interfaces to control in *realtime* various digital sound processes.

Concerning fundamental studies, the first detailed measurements of bowing parameters were performed by Askenfelt [1, 2], using a bow equipped with sensors for measuring bow force, bow position and bow-bridge distance. Knowledge of how players make use of these parameters provide an important key to violin performance, which could for example be useful for controlling physical models of bowed string instruments or applications in music education.

The *Hyperbow* [6], or more recently the *Augmented violin* [3, 5] are two examples, among others, of modified bows used in live performance. In these particular cases, accelerometers are placed on the bow. The acceleration signal can be used to detect bow stroke and in some cases to recognize bowing styles [5].

Bow velocity is an important parameter to characterize bowing, and is one of the most important input parameters for playing a physical model of the violin. Several acoustic studies ([1, 2]) have also clearly shown the relationship between velocity and sound quality. Nevertheless, bowing velocity can be difficult to measure accurately in a playing situation. Velocity can potentially be derived from video tracking or accelerometer signals. Difficulties arise in both cases:

- Computing accurate velocity profile from video tracking system generally requires the reconstruction of the bowing orientation in 3D space. Moreover, expensive camera for high temporal resolution is generally necessary for fast movements such as bowing attacks. The use of systems such as 3D motion capture system is generally limited to the laboratory environment.

- Computing velocity over a longer time span by integration of the accelerometer signal is problematic. For example, the accelerometer used in the augmented bow [3, 5] is sensitive to both inclination and acceleration (generally referred to as static and dynamic acceleration), which means that there is a variable amount of drift present in the integrated signal.

The method described here, combining the use of both an accelerometer and standard video, allow us to overcome such problems. In particular, we present a method compensating for this drift, enabling therefore the reconstruction of bow velocity from the bow acceleration signal. First, we explain the general principles of the method (section 2). In section 3 the method is assessed using data obtained with a 3D motion capture system, to evaluate the potentiality and limitations of the method. The results are discussed in section 4.

2. BOW VELOCITY RECONSTRUCTION

We first explain the general principle of the setup and reconstruction method. Second, we describe a particular pilot study that serves as a proof-of-principle example.

2.1 Setup

The setup is shown in Fig.1. The two main components are a fixed camera and an accelerometer placed on the bow. The accelerometer axis is set to be parallel to the bow axis. Standard video processing techniques are used to track the bow movements. For example, two color markers can be placed on the bow enabling robust tracking. The accelerometers can be part of the wireless modules described in reference [5] and [6]. Note that simpler implementation is also possible using wired connection between the accelerometer and a sensor A/D interface.

Two angles in this setup are important to consider, both varying significantly during the playing. First, α is the angle between the bow and the vertical direction. The variation of α is the main responsible for the drift in the acceleration signal. Second, β is the angle between the axis of the camera and the direction perpendicular to the bow. The relation between the velocity along the length-axis of the bow and the velocity observed by the camera is then characterized by a scaling factor $\cos(\beta)$.

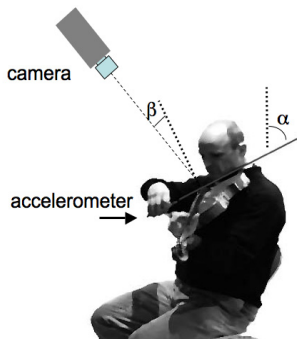


Fig. 1. Setup.

2.2 Computation

As already mentioned, velocity computed from the accelerometer signal typically contains a drift over time. The reconstruction method allows for the compensation of this drift in the integrated signal, by detecting moments when bow direction change. At such moments the bow velocity is equal to zero (referred below as “breakpoints”).

Using such information, the accelerometer signal is piece-wise integrated between these “breakpoints”. After this step, the velocity curve shows discontinuities at the “breakpoints”, which can be removed making an assumption on the form of the acceleration drift. The simplest assumption corresponds to a constant drift in the acceleration signal between two breakpoints, corresponding to a constant α during a bow stroke. This constant drift in the acceleration produces piece-wise constant slopes in the integrated signal. Such linear trends can

be simply computed and removed, resulting in a velocity profile that is continuous at the breakpoints.

To obtain the zero-crossings of the bow velocity, a simple video camera can be sufficient. Key points on the bow and the violin are tracked using video processing techniques in order to detect the moments when the bow changes direction. Errors induced by the camera position and the low frame rate of standard video camera are addressed in section 3.

The reconstructed velocity can potentially have a high spatial and temporal resolution (depending on the accelerometer and on A/D conversion system). Therefore, the method offers an easy-to-implement and a cost-effective alternative to expensive commercial motion capture systems to obtain bow velocity signals in violin playing.

The next section describes the implementation of the reconstruction method in a pilot experiment, using a normal video camera and the augmented bow.

2.3 Pilot experiment

A pilot experiment was performed to test the feasibility of the velocity reconstruction method. The bow acceleration was measured with the augmented bow, developed at IRCAM [3, 5]. The bow was equipped with two Analog Device ADXL202 acceleration sensors at the frog, and the acceleration data was sent wirelessly to a RF receiver, connected to a sensor acquisition system, Ethersense [4]. The acceleration data was digitized on 16 bits at the frame rate of 500 Hz.

The video data was obtained with a Sony digital handcam (type DCR-TRV245E). For the measurements reported in Figure 2, two points were marked using differently colored pieces of fabric: one attached to the curl of the violin and the other to the wrist of the player’s bowing arm. The camera was positioned in front on the right side of the player, so that the marked points were visible, and the bow motion could be clearly observed (the influence of such setup configuration will be discussed in section 3). The color markers were tracked using Eyesweb software [9], by selecting the pixels with the specified colors and calculating the centre of gravity of the observed pixel regions. The frame rate of the video data was 25 Hz.

The acceleration and video data were synchronized by aligning two synchronization events at the beginning and the end of the recording. The synchronization features were obtained by

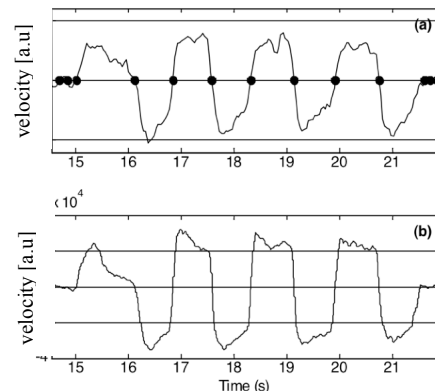


Fig. 2. Détaché bowing on one string. Velocity signal derived from video data (a). The detected zero-crossings are indicated by dots. Reconstructed velocity signal (b) obtained by piece-wise integration of the acceleration of the bow.

tapping with the bow on the curl of the violin, which left a measurable trace in both data series.

From the video data, the position of the wrist marker was obtained in the camera reference. The motion of the violin marker was subtracted to obtain the relative motion of the wrist marker. The derivative was computed and smoothed using Savitzky-Golay filtering (order 5, frame size 11). The zero-crossings of this velocity signal were detected and refined using linear interpolation to obtain a more precise time estimate. Figure 2a shows a series of recorded *détaché* bowing styles: a single tone was repeatedly played to avoid string crossings.

In the next step, the velocity was reconstructed from the acceleration signal of the bow, shown in Figure 2b. The reconstructed velocity profile is coherent as there is an equal repartition between positive and negative velocity, which is in accordance with the performed movement. The reconstructed velocity data has a temporal resolution equal to the accelerometer sampling rate, i.e. 500 Hz. In comparison, the sampling rate of the velocity profile from the video data is only 25 Hz. Moreover, Figure 2 shows that the velocity from the video data is significantly noisier than the velocity profile reconstructed from the accelerometer.

3. METHOD ASSESSMENT

The velocity reconstruction method was further evaluated using 3D motion capture data. During this experiment the bow acceleration was measured simultaneously to the motion of the bow and the violin. The motion capture data provided a reliable reference signal of the actual bow velocity, which was used for quantitative comparison with the reconstructed bow velocity.

The evaluation of the method was performed on three aspects. First, velocity reconstruction is evaluated quantitatively, using an optimal velocity signal. Second, the influence of the video camera position is addressed. Last, the influence of video latency and jitter is quantified.

3.1 Motion capture setup

The motion of the violin and the bow was tracked using a Vicon MX system, with 12 cameras at the frame rate of 150 Hz. The motion capture data was smoothed using Savitzky-Golay filtering (order 3, frame size 9).

Bow acceleration was measured with the same augmented bow described in section 2.3. The sample rate of the acceleration data was 500 Hz. Acceleration data was smoothed using Savitzky-Golay filtering (order 3, frame size 25 – corresponding with the mocap smoothing parameters in the time domain). The data selected for the evaluation was a recording of scales with different bowing styles: *détaché*, *martelé* and *spiccato* played at 60 bpm.

3.2 Velocity reconstruction validation

We first validated the velocity reconstruction method in the case where the velocity zero-crossings were accurately known from the motion capture data, in order to show the achievement of the method under optimal conditions. As the reference signal the velocity along the length-axis of the bow was taken, which corresponds with the actual bow velocity at the string. The reference signal was computed from the motion capture data, compensating for the motion of the violin. During the first validation step the zero-crossings of the reference velocity were used as breakpoints for the reconstruction.

Figure 3 (top) shows the acceleration signal of the bow during a *détaché* scale. It clearly demonstrates the influence of angle α

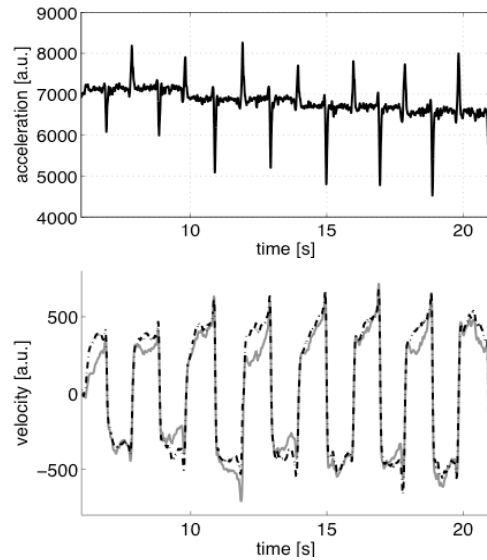


Fig. 3. (top): accelerometer signal during a *détaché* bowing. (bottom): reference velocity (dotted) and reconstructed velocity (plain)

on the acceleration signal, the different offset levels corresponding to playing on different strings.

Figure 3 (bottom) shows the reconstructed bow velocity, as well as the bow velocity reference for a series of *détaché* bow strokes using the reference signals zero-crossings as breakpoints. It shows that the reconstructed velocity is in relatively good agreement with the reference velocity. Note that similar results were also found for the other bow stroke types. For example, the correlation coefficients between the reference and the estimated velocities are 0.984 for *détaché*, 0.998 for *martelé* and 0.987 *spiccato* (computed on ascending and descending scale at 60 bpm). These high correlation values demonstrate the validity of the method, when the velocity zeros are accurately determined.

3.3 Influence of the camera viewpoint

The video camera implies a 2D projection of the markers movement. Such a 2D projection is in most cases sufficient since we are not interested in the actual velocity profile from the video data, but only in the determination of the zero-crossing. The most important point is thus to guarantee a sufficiently a high resolution image of the bow movements. Poor resolution can lead to important errors in the determination of the velocity zero-crossing. The optimal viewpoint is therefore the one providing the best overall resolution of the bow displacement.

As mentioned in section 2.1, the projected bow velocity depends on the angle of the camera with the bowing direction with a scaling factor of $\cos(\beta)$. For an optimal projection the video camera should therefore be placed perpendicular to the length-axis of the bow, for example above the player (see Fig. 1).

3.4 Influence of video latency and jitter

Latency and jitter between the accelerometer and video streams could occur, especially if no synchronization mechanism is operated. Such phenomena will mainly affect the timing of detected zero-crossings. In this section, we evaluate the effect of altering such timing in the reconstruction velocity.

An increasing latency was added to the reference breakpoints. The latency varied from 0 to 80 ms by steps of 20 ms. An

uniformly distributed random number of span 40 ms was added to simulate a jitter effect. The consequence on the correlation coefficient is shown in Figure 4.

Détaché showed a better robustness to jitter than the two others with a smaller variation. Nevertheless, when the latency is kept sufficiently low, the jitter is not a major source of error for the three tested bowing styles.

For *détaché*, the correlation coefficient decreased from 0.984 to 0.87 when latency increases to 40 ms. However, the correlation coefficient stabilized as the latency increases to 80 ms, dropping only to 0.84. For *martelé*, the method is less sensitive to small latencies. The correlation coefficient is higher than 0.85 for latencies less than 60 ms. *Spiccato* appeared to be the more sensitive to latency/jitter effects as the correlation coefficient fell almost linearly.

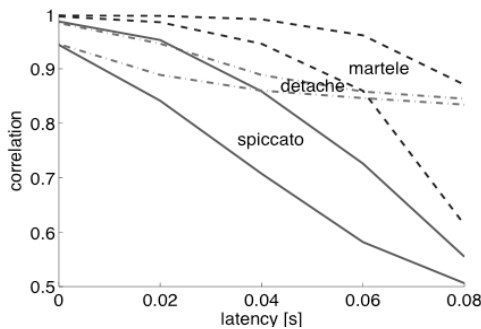


Fig. 4. Correlation factor according to latency and jitter for the three bowing styles. For each bowing style, the upper line is without simulated jitter and lower line with simulated jitter (40 ms).

3.5 Influence of the movement of the player

Except for the errors due to projection and synchronization issues, there are two other possible error sources associated with the reconstruction method. First, as already mentioned it was assumed that the drift between the breakpoints is constant, implying that the angle α in Fig.1 does not change. However, this assumption is not necessary valid in any playing situation, as the player can vary the angle of the bow, especially on the outer G and E strings. Second, the acceleration is measured relative to a fixed reference, rather than the moving violin. This means that movements of the player not directly related to playing are contributing to the reconstructed velocity as well. Thus, the achievement of the method could be dependent on the complexity of the bowing pattern, as well as the amount of additional (expressive) movements by the player.

4. DISCUSSION AND CONCLUSIONS

The overall results show that the method has good potential to reconstruct velocity profile with high temporal resolution. The method is simple to implement and cost-effective compared high-performance motion capture system.

This method could be useful for fundamental studies of bowing gesture, in particular cases where expensive motion capture system systems are not available or too cumbersome to use. The results shown here are promising, but further validation is required to fully characterize the precision and accuracy of this method.

The reconstruction method could also be used in live performance. Nevertheless, the implementation describe here cannot be used as a “strict” realtime system, in the sense that the reconstruction method implies an inherent variable delay.

As a matter of fact, the accelerometer drift is not corrected continuously but at discrete time (i.e. velocity zero-crossing). Moreover, as shown in section 3.4, accurate results might require synchronizing the accelerometer and video streams, which would add an additional delay.

Nevertheless, the system can still be useful in performance situation, where detailed information of the bowing gesture is desired and a delayed response is manageable. In such cases, the reconstructed velocity profile can reveal to be very helpful, since the bow velocity is one of the fundamental parameter in the bowing gesture.

The method we described here could also be very valuable for pedagogical applications. For example, accurate information on the playing regularity of specific bow strokes could provide the students with helpful information. In such cases, the information is needed only after the playing.

5. ACKNOWLEDGEMENTS

This work was supported by the Swedish Research Council, project: Bowed String Analysis (621-2001-2537) and the Cost287-ConGas action. We are grateful to Stephane Donikian (IRISA) and Franck Multon (Université Rennes 2) for giving us access to their motion capture system and their precious help during the mocap sessions. We also thank Christophe Desjardins and Emmanuel Fléty for their helpful participations in this project.

6. REFERENCES

- [1] Askenfelt, A. *Measurement of bow motion and bow force in violin playing*. J. Acoust. Soc. Am. 80 (4), pp. 1007-1015, 1986.
- [2] Askenfelt, A. *Measurement of the bowing parameters in violin playing. II: Bow-bridge distance, dynamic range, and limits of bow force*. J. Acoust. Soc. Am. 86 (2), pp. 503-516, 1989.
- [3] Bevilacqua, F., Rasamimanana, N., Fléty, E., Lemouton, S. & Baschet, F. The augmented violin project: case studies, composition and performance report. In *Proceedings of the Conference on New Instruments for Musical Expression*, NIME 2006.
- [4] Fléty, E., Leroy, N., Ravarini, J. & Bevilacqua, F. Versatile sensor acquisition system utilizing network technology. In *Proceedings of the Conference on New Instruments for Musical Expression*, NIME, 2004
- [5] Rasamimanana, N., Fléty, E., Bevilacqua, F. *Gesture Analysis of Violin Bow Strokes*. GW 2005, Lecture Notes in Artificial Intelligence 3881, S. Gibet, N. Courty, and J.-F. Kamp (Eds.), pp. 145–155, 2006.
- [6] Young, D., The hyperbow controller: real-time dynamics measurement of violin performance. In *Proceedings of the 2002 conference on NewInterfaces for Musical Expression*, NIME, Dublin, 2002
- [7] Foxlin, E., Naimark, L., "VIS-Tracker: A Wearable Vision-Inertial Self-Tracker". In IEEE Conference on Virtual Reality (VR 2003). Los Angeles, CA.
- [8] Kwon, D. Y., Gross, M., Combining Body Sensors and Visual Sensors for Motion Training, *Proceedings of ACM SIGCHI ACE 2005*, pp. 94-101.
- [9] <http://www.eyesweb.org>

ARTICLE NIME 2006B

THE AUGMENTED VIOLIN PROJECT :
RESEARCH, COMPOSITION AND PERFORMANCE REPORT.

F. BEVILACQUA, N. RASAMIMANANA, E. FLÉTY, S. LEMOUTON AND
FLORENCE BASCHET

The augmented violin project: research, composition and performance report

Frédéric Bevilacqua, Nicolas Rasamimanana, Emmanuel Fléty, Serge Lemouton and
Florence Baschet

IRCAM - Centre Pompidou
CNRS STMS UMR 9912
Gesture working group
1 place Igor Stravinsky
75004 Paris – France
+ 33 1 44 78 48 31

{frederic.bevilacqua, nicolas.rasamimanana, emmanuel.flety, lemouton}@ircam.fr, fbaschet@noos.fr

ABSTRACT

In this paper we present the augmented violin developed at IRCAM. This instrument is an acoustic violin with added sensing capabilities to measure the bow acceleration in realtime. We explain first the approach we developed to characterize bowing styles. Second, we describe the realtime implementation of the bowing style recognition system. Finally we describe an electro-acoustic music composition, *Bogenlied*, written for the augmented violin.

Keywords

Augmented violin, hyper-instrument, bowing styles, mapping.

1. INTRODUCTION

The augmented violin project started in 2003 at IRCAM following the interest of several composers to use violin gestures for the control of electronic processes. This project triggered the creation of an interdisciplinary working group on gesture analysis and musical interfaces. One research goal of this group is to work on the concept of “augmented instruments”, i.e. acoustic instruments with added gesture sensing capabilities, which is similar to the *hyperinstruments* pioneered at MIT. We believe that such an approach is particularly fruitful for both fundamental gesture research and artistic endeavors.

Different approaches are possible with “augmented instruments”. First, sensors can be utilized to add control possibilities that are not directly related to normal playing techniques. For example, various buttons can be added to the body of the instruments. In such a case, the use of sensors implies new gestures for the player. Second, sensors can be applied to capture normal playing gestures. This paper is related to this type of approach, which poses a fundamental question: to what extent can an instrumental gesture, mastered in a particular context, be used in another (or larger) context of musical expression? Such a questioning is actually very fruitful,

and remains valid for any types of musical interfaces.

This paper focuses on the ongoing “augmented violin” project. We report here first a study on bowing styles characterization, and the implementation of a system performing realtime bowing style recognition. Second, we describe a music composition, *BogenLied* by Florence Baschet¹ written for the augmented violin and utilizing the analysis system.

2. RELATED WORKS

Different systems have been developed to directly or indirectly measure the violin “gesture parameters”. Several systems have been developed over the years using various sensing systems [4][10][9][3]. An alternative approach is to use audio features as a trace of the performed gestures [6][2].

In most of the works, gesture data has been used as a direct control of sound filters, or as the input for physical model synthesis [8]. Fewer works have reported on interpreting the gesture data, in order to provide high-level parameters used in the mapping design. Bernd Schoner [7] used the gesture input to statistically estimate the corresponding sound features in order to drive more expressive synthesis. Chad Peiper [5] used decision trees to classify different bowing styles.

3. APPROACH

Augmented instruments are based on traditional instruments, which gestures are *a priori* defined. In the case of the violin, the different types of bow strokes form a widely accepted and formalized set of gestures. Composition for strings includes bow strokes indications such as *détaché*, *martelé*, *ponticello*, etc. On a finer level, the mastering of various bowing articulations is part of the player’s skill.

Our approach has been to use this gesture “vocabulary” as a starting point to build the interaction between the player and the electronics. This is somewhat similar to the approaches of Schoner or Peiper, since we propose to build an “interpretation level” from the data stream, in order to facilitate the mapping between gestures and sounds.

Our first goal was hence to study the relationships between bowing styles and the various sensors data. Such studies allowed us then to build real-time systems that can “interpret”

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
NIME 06, June 4-8, 2006, Paris, France.

Copyright remains with the author(s).

¹ Performed at Dijon, France, Nov. 26 2005 Whynote Festival (Premiere), Violin: Anne Mercier (Ensemble Itinéraire); Jan. 07 2006 Maison de la Radio, Paris, France.

the low-level captured parameters into high-level parameters related to bowing styles.

This approach was found in resonance with the compositional approach of Florence Baschet: “With *BogenLied*, my aim as a composer was to focus my attention to the fine instrumentalist gestures, and to use such gesture articulations as the input of an interactive system. Precisely, *BogenLied* is an attempt to create a mixed –acoustic and electronic– sound space, resulting from a sensitive interactive relationship between the soloist and electro-acoustic system. Such interaction would be ideally similar to the type of close relationship between two musicians, as typically found in chamber music”.

4. THE AUGMENTED VIOLIN

Two prototypes have been successively built. They are described below.

4.1 Prototype I

Our first prototype of the wireless bow measurement system is composed of a small electronic board with a microcontroller, two ADXL202 accelerometers from Analog Devices and a digital radio transmitter (fig.1). The principle of the measurements is similar to the techniques developed by Joe Paradiso and Diana Young [4][10]. Our first prototype utilizes a special radio transmitter that enables collision detection on the carrier and therefore permits to share the transmission bandwidth with other bows equipped similarly. The bow module was very satisfactory. The only drawback is its relative big size due to the thickness of the radio transmitter and the batteries.

4.1.1 Bow acceleration

Accelerations are not analogously sampled but measured using PWM capture with accurate counters and timers. This technique delivers acceleration values on a range of 26000 points, which is significantly better than the use of the internal 10 bit ADC of the microcontroller. A wireless receiver is placed on an Ethersense [1] daughter board. The Ethersense sends the digital data from the accelerometers (as well as others signals) to a host computer through Open Sound Control (OSC).

4.1.2 Bow position

The bow position extraction is directly inspired from the electric field sensing measurement described by Paradiso: two different signals are emitted from the bow tip and frog. A capacitive coupling plate placed behind the bridge collects the mix of the two signals, “tip” and “frog”, which relative intensities depend on the bow position [4]. An Ethersense [1] daughter board demodulates the mix of bow positioning signals. The material used to make the resistive strip placed on the bow was taken from a S-VHS tape. This tape features a homogenously distributed electrical resistance, over its whole length. Two strips of the tape were glued one on the other to protect the resistive side and to make the overall resistance in a more adequate range.

A software calibration was developed to compute the bow position based on the measurement of the two signals “tip” and “frog”. The calibration was based on careful measurements for a large set of different bow positions. However, our effort to perform accurate and reliable bow position measurement was deceptive. The main difficulty with this technique is due to the fact that the two measured signals are not sufficient to determine without ambiguity the bow position. This problem especially occurs at the bow extremities where the surface of the “bow plate” diminishes. In addition, the right hand causes a drastic modification of the signals as the impedance of the body

interferes with the system, causing a decrease in the signals intensities. An additional difficulty comes from bowing techniques implying significant variations of bow angle, which affects the coupling capacity between the plates.

Overall, the distance measurement was found useful only for qualitative measurement, but problematic for accurate position measurement. Other methods for position measurement are currently experimented.

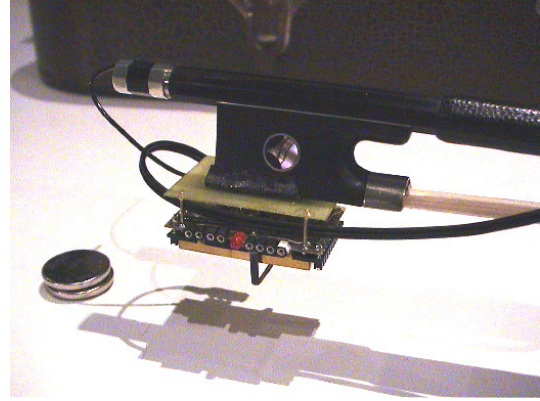


Figure 1: Prototype I.

4.2 Prototype II

A second prototype was built using a subminiature (but single frequency) radio transmitter. The custom-made batteries holder and the second accelerometer were mounted on an extension of the PCB placed on the left side of the bow frog (fig. 2). This second prototype is smaller and lighter than the prototype I (total weight of 17g). The overall thickness of this device was significantly reduced, limiting therefore the risk of scratching the violin. Soft foam was added to totally suppress any possibility of such an accident (not shown on fig 2). Moreover, it consumes less current, allowing 1h30 of continuous playing.

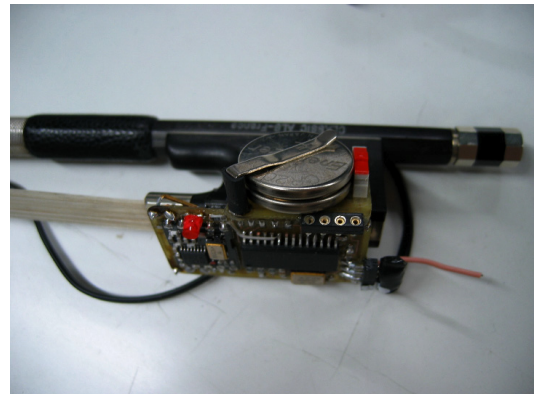


Figure 2: prototype II

5. BOW STROKE STUDIES

We focused our first studies on the following bowing styles: *détaché*, *martelé* and *spiccato*. After recording these bow strokes in various musical contexts (scales, musical phrases) and with different players, a complete offline analysis has been performed. This analysis has been reported in reference [11] and we summarize here only the important points.

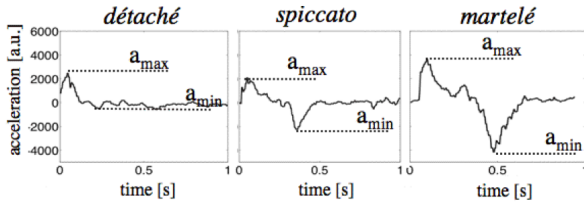


Figure 3. Typical acceleration curves for *détaché*, *martelé* and *spiccato* (from left to right respectively)

- The acceleration curve in the bowing direction shows typically two principal peaks, one positive, a_{max} and the second one negative, a_{min} (for upbow, see fig.3). The *détaché* stroke can be seen as a degenerated case where the second peak is close to zero.

- The use of two parameters a_{max} and a_{min} allows for the clustering of these bowing styles. Recognition rate using the standard *knn* method (k-nearest neighbors) leads to high recognition rates.

- The parameter $(a_{max})^2 + (a_{min})^2$ can be interpreted as a bow stroke “intensity”. This parameter is correlated to the energy given to the bow by the musician. It is also correlated to dynamics when the bow is in contact with the strings. For bouncing bowing styles, such as *spiccato*, this parameter gives an indication of the gesture “intensity” rather than the loudness of the sound.

The first study was performed with two violin players, amateur and professional, and showed very consistent results. A second study was performed with a class of 12 students. The results demonstrated that the characterization we proposed remained valid over this larger set of players: the three bow strokes can still be clustered with a_{min} and a_{max} . However, idiosyncratic behaviors were also found, showing that a universal calibration might not be reliable. Interestingly, these results also showed us that this type of analysis could potentially be useful for pedagogical applications. For example, the bowing characterization could provide the students with a complementary feedback to improve bow regularity and/or dynamics. Such potential applications for pedagogy are currently studied.

6. REALTIME BOWSTROKE ANALYSIS

A real time implementation of the bow stroke characterization described in the previous section was implemented in Max/MSP, using several objects of the library MnM [12], dedicated to gesture analysis. The different steps of the gesture analysis are described in Fig 4.

A median filter is used on the x-acceleration data (main bow axis). The dataflow is then split into two different processes run in parallel.

First, a parameter related to the “intensity” of the bowing is computed from the acceleration curve. Precisely, the maximum of the absolute value of the acceleration, computed on a sliding window, is output.

Second, a distinct process allows for the segmentation and characterization of the bow strokes. The varying baseline due to the coupling between angle and acceleration is first removed, (assuming a linear offset). The segmentation is performed then in two steps. First, accelerometer peaks are determined on a sliding window. Second, a procedure sorts and labels the various peaks. One of the difficulties resides in the fact that the “maximum” acceleration a_{max} alternates between bow changes:

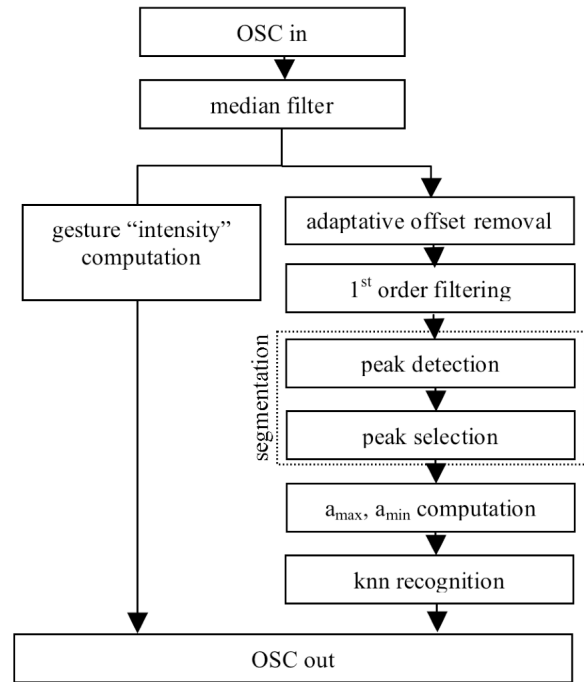


Figure 4. Schematic of the realtime gesture processing

in upbows a_{max} appears as a local maximum in the raw signal while it appears as a local minimum for downbows. Our selection and labeling procedure is based on the timing between the different peaks. Precisely, we used the fact that the time difference between a_{max} and a_{min} is typically smaller than 100 ms.

This segmenting procedure is fundamentally underdetermined for a sequence of fast notes, such as a tremolo (when using the accelerometer signal only). In such a case, the peak labeling procedure is ambiguous; different bow strokes might give rise to very similar acceleration curves. Such a problem could be solved by the combination of both audio and acceleration data.

Once the segmentation is performed, the associated a_{max} and a_{min} parameters of each bow stroke are computed. Fig 5 shows an example (as displayed in Max/MSP) of clustered points obtained in the a_{max} and a_{min} plane, with *détaché*, *martelé* and *spiccato* (playing scales).

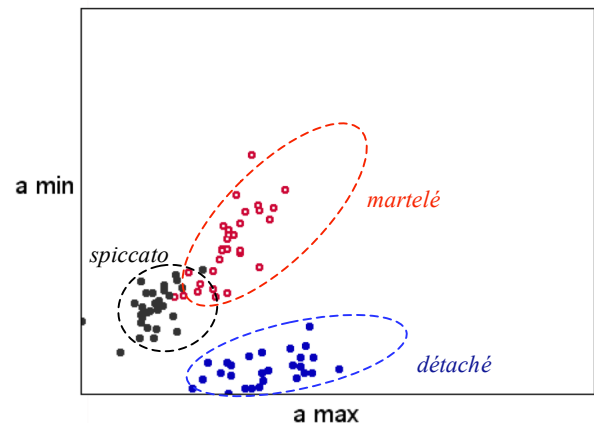


Figure 5. LCD object from Max/MSP showing the clustering on the three bow strokes (real-time): *détaché* (blue), *martelé* (red), *spiccato* (black).

Finally, a *knn* algorithm (k-nearest neighbors) is run to classify each bow stroke as *détaché*, *martelé* or *spiccato*, according to a previously recorded bow strokes database. Weights related to bowing styles are computed using the number of nearest neighbors of each class. The weight averages over several notes are also computed.

The final results, i.e. gesture intensity and bow stroke weights are sent through Open Sound Control (OSC) to the computer controlling the sound processes.

7. COMPOSITION WORK: *BOGENLIED*

7.1 Setup

The diagram (figure 6) shows the configuration used in the performance setting. Two separate computers (two Macintosh G4 PowerBooks) are running Max/MSP, and communicating using OSC. The “sound processing” computer generates the digital sound environment using the live violin sound, captured by microphone. Various parameters of the sound processing are controlled by the gesture data transmitted from the “gesture-processing” computer. The electronic sound is spatialized (using the “Ircam spat”) and rendered by a hexaphonic sound diffusion system.

7.2 *Bogenlied* form.

As shown in fig. 7, the musical form of *BogenLied* is a simple

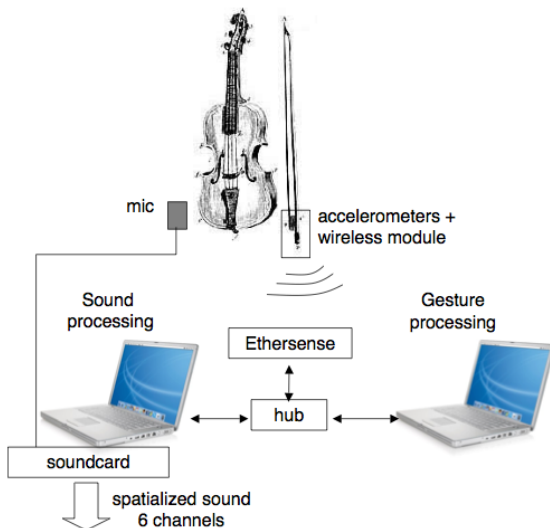


Figure 6. Setup for the piece *Bogenlied*

linear form divided into nine sections, alternatively with and without electronics. In section *II*, *IV* and *VI*, the musical writing is focused on specific articulations of bow strokes: *détaché*, *martelé*, *spiccato*. Each of these sections is associated to specific compositional materials, as well as to specific electronic sound processes (described in section 7.3).

In section *VIII*, the three bowing styles appear successively, along with their associated compositional material. In this section, the choice of the electronic processes is driven by the recognition system.

In last section (*IX*), the performer plays hybrid bow strokes, which sequence appears as a series of “mutations” from one articulation to another one. As described further below, the electronic sound is then built as a combination of the electronic processes performed in section *II*, *IV* and *VI*.

section	I	II	III	IV	V	VI	VII	VIII	IX
bowing style		<i>détaché</i>		<i>martelé</i>		<i>spiccato</i>		<i>all 3 bowing styles</i>	<i>hybrid</i>
sound	acoustic	mixed	acoustic	mixed	acoustic	mixed	acoustic	mixed	mixed
Gesture analysis		gesture intensity		gesture intensity		gesture intensity		gesture intensity + bow stroke recognition	gesture intensity + bow stroke recognition

Figure 7. Structure of *Bogenlied*

7.3 Audio processing

The Max/MSP sound patch contains two granular synthesis modules and several effects. The first granulator is used to produce drones (pedal-notes) triggered by the bow strokes. A separate drone is linked to each bowing styles: *G*, *A*, and *E* for *détaché*, *martelé* and *spiccato*, respectively.

The second granulator is used for real-time processing of the violin sound. The audio signal from the microphone is stored in a 5 second circular buffer, and 100 to 200 ms grains are played back from this buffer. The granular synthesis sound is processed through several standard effects: frequency-shifter, harmonizer, filter and vocoder. Different combinations and parameterizations of these effects are used in each section. Some parameters are controlled in realtime by the gesture data, as explained next.

7.4 Gesture Mapping

The mapping between the gesture data and sound processing parameters is built during the piece as a gradual superimposition of three different mapping modalities:

- Continuous* mapping, applied indifferently to all bow strokes
 - sections *II*, *VI*: the gesture “intensity” is linearly mapped to the grain density (number of grains played simultaneously)
 - section *IV*: the gesture “intensity” is mapped to the grain transposition spread (small gestures induce small or no transposition, strong gestures cause all the grains to be randomly transposed around the played pitch).
- Selective* mapping using bow stroke recognition:
 - sections *VIII* and *IX*: bow stroke recognition is used to select the type of sound processing. The recognition of a *détaché*, *martelé* or *spiccato* recalls the presets of section *II*, *IV* or *VI*, respectively (including the corresponding drones).
- Mapping *mixing*, using bow stroke characterization (of hybrid bow strokes):

- section IX, the performer plays hybrid bow strokes. The gesture analysis computes for a given bowstroke three weights, corresponding to “likelihood” to be related to the different bowing styles (in other words a given bow stroke is considered as a mix between different bowing styles). These weights are then used to control the mixing levels of the drones associated with each bowing style. Each bow stroke is thus colored by the combination of these three pitches reflecting continuously the quality of the articulation.

8. DISCUSSION AND PERSPECTIVES

We report here developments related to our current “augmented violin” project. From the technological point of view, our prototypes were tested on several experiments and performances and found to be robust. The added weight on the bow seems manageable for the players: the various professional violinists we worked with agreed to play with such a constraint.

We described the real-time implementation of a bow stroke analysis framework that was reported previously [11]. This approach requires the segmentation of the acceleration data stream in separate bow strokes. Such a task is difficult when the acceleration signal alone is used, due to various artifacts. The algorithm we designed was generally satisfactory, but was not applicable to fast notes, such as tremolo. We are currently investigating other approaches, based for example on Hidden Markov Models, in order to overcome such limitations.

BogenLied is the first piece written for the IRCAM augmented violin. It is worth to note that this piece (composed by Florence Baschet) has been developed in the context of an interdisciplinary workgroup. The piece takes advantage of the gesture analysis we described. In particular, the mapping is expressed from elements of the musical language, in order to create an electronic environment that performers can apprehend intuitively. Our collaboration with the violinist Anne Mercier confirmed us that such an approach is very effective.

The real-time analysis was reliably used in performances of *BogenLied*. Overall, the recognition system was satisfactory. In particular, the performer was able to control the mapping easily with her own gestures, and she could feel that the system was reacting well to her gesture. A promising point was the possibility to characterize “hybrid” bow strokes by continuous parameters. This feature seems to offer pertinent information in complex musical phrases.

In the near future, other performers will experiment with this system, and we are expecting interesting comparisons. In particular, the modules will be used with other string instruments such as the viola, cello and double bass. This collaborative work has provoked a high interest at IRCAM and several other artistic works are currently in progress.

9. ACKNOWLEDGMENTS

We would like to thank Anne Mercier, Alain Terrier and Jean-Loup Graton for their precious contribution and help.

The I-MAESTRO project is partially supported by the European Community under the Information Society Technologies (IST) priority of the 6th Framework Programme for R&D (IST-026883, www.i-maestro.net, www.i-maestro.org). Thanks to all I-MAESTRO project partners and participants, for their interests, contributions and collaborations.

10. REFERENCES

- [1] E. Fléty, N. Leroy, J.-C. Ravarini, and F. Bevilacqua. Versatile sensor acquisition system utilizing network technology. In *Proceedings of the Conference on New Instruments for Musical Expression, NIME*, 2004.
- [2] C. Goudeseune, G. Garnett, and T. Johnson. An interface for real-time classification of articulations produced by violin bowing. In *Proceedings of the Conference on Human Factors in Computing Systems*, 2001.
- [3] D. Overholt. The overtone violin. In *Proceedings of the Conference on New Instruments for Musical Expression, NIME*, 2005.
- [4] J. A. Paradiso and N. Gershenfeld. Musical applications of electric field sensing. *Computer Music Journal*, 21(2), 1997.
- [5] C. Peiper, D. Warden, and G. Garnett. An interface for real-time classification of articulations produced by violin bowing. In *Proceedings of the Conference on New Instruments for Musical Expression, NIME*, 2003.
- [6] C. Poepel. Synthesized strings for string players. In *Proceedings of the Conference on New Instruments for Musical Expression, NIME*, 2004.
- [7] B. Schoner, C. Cooper, C. Douglas, and N. Gershenfeld. Cluster-weighted sampling for synthesis and cross-synthesis of violin family instrument. In *Proceedings of the International Computer Music Conference, ICMC*, 2000.
- [8] S. Serafin and D. Young. Bowed string physical model validation through use of a bow controller and examination of bow strokes. In *Proceedings of the Stockholm Music Acoustics Conference, SMAC*, 2003.
- [9] D. Trueman and P. R. Cook. Bossa: The deconstructed violin reconstructed. In *Proceedings of the International Computer Music Conference, ICMC*, 1999.
- [10] D. Young. The hyperbow controller: Real-time dynamics measurement of violin performance. In *Proceedings of the Conference on New Instruments for Musical Expression, NIME*, Dublin, Ireland, 2002.
- [11] N. Rasamimanana, E. Fléty, F. Bevilacqua, Gesture Analysis of Violin Bow Strokes, GW 2005, Lecture Notes in Artificial Intelligence 3881, S. Gibet, N. Courty, and J.-F. Kamp (Eds.): pp. 145–155, 2006.
- [12] F. Bevilacqua, R. Muller, N. Schnell, MnM: a Max/MSP mapping toolbox. In *Proceedings of the 2005 International Conference on New Interfaces for Musical Expression, NIME*, Vancouver, p. 85–88, 2005.

