

Découvrir le traitement du signal et de l'image grâce à des projets personnalisés : du traitement de la parole à la télécommande virtuelle

A. Giremus², R. Mégret¹, E. Grivel¹, O. Lavalie³ et Y. Berthoumieu¹

¹ENSEIRB, 1 av. du Dr. A. Schweitzer, B.P. 99, F-33402 Talence Cedex

²Université Bordeaux 1, UFR de Physique, Dpt. EEA, Cours de la libération, F-33405 Talence Cedex

³ENITAB, 1 cours du Général de Gaulle, F-33175 Gradignan Cedex

RESUME : Pour sensibiliser les étudiants au traitement du signal et de l'image, nous avons mis en place des projets personnalisés par groupe de quatre étudiants sur un semestre, à raison d'une demi-journée par semaine. Les sujets proposés par l'équipe pédagogique permettent de couvrir une large gamme d'applications (de l'audio aux systèmes de communications mobiles, de l'image à la vidéo, etc.) et peuvent reposer sur des problématiques industrielles. Ce projet vise à décloisonner les enseignements, à sensibiliser les étudiants à nos thématiques, à les rendre davantage acteurs de leur formation et à leur permettre d'avoir un avant-goût du type d'action qu'ils pourront mener en entreprise. Réalisé en équipe, il constitue un cadre idéal pour confronter l'étudiant à la recherche documentaire et à la gestion de projet. Il permet en outre d'intégrer des enseignements non scientifiques comme l'anglais (au travers de la lecture des documents, la rédaction d'un rapport et d'une soutenance en anglais qui donnent lieu à une évaluation intégrée dans la note finale). Dans cet article, nous illustrons notre démarche pédagogique au travers de trois exemples : le premier porte sur le traitement de la parole, le deuxième sur le traitement d'image et le troisième sur le traitement vidéo.

Mots clés : traitement du signal, image, vidéo, projet.

1 INTRODUCTION

L'enseignement du traitement du signal et de l'image (TdSI) au niveau BAC+4 ou BAC+5, en école d'ingénieurs ou à l'université au niveau Master, repose généralement sur des enseignements classiques organisés en cours, TD et TP et éventuellement des projets d'initiation. Cette structuration permet aux étudiants de se familiariser aux concepts du TdSI et de manipuler des notions de base de la théorie du signal. Pour sensibiliser les étudiants à nos thématiques, les rendre davantage acteur de leur formation et leur permettre d'avoir un avant goût du travail qu'ils pourraient mener tout d'abord en stage de fin d'étude puis en entreprise, d'autres approches pédagogiques doivent être envisagées.

Ainsi, depuis quelques années, à l'ENSEIRB, nous faisons l'expérience d'une forme pédagogique complémentaire que constituent les projets personnalisés. Il s'agit de projets par groupes de quatre étudiants, encadrés par un enseignant-chercheur sur un thème qui favorise le décloisonnement de plusieurs enseignements déjà suivis par les étudiants tant en traitement du signal et des images, en programmation (Matlab, C, Java, html, etc.) et/ou en implantation numérique (VHDL, Système C), qu'en anglais et gestion de projet.

Ce projet est formateur tant dans son contenu technique que dans son organisation. L'équipe pédagogique se présente comme un client exprimant des besoins auxquels doivent répondre le groupe de quatre étudiants. Le projet étant planifié sur un semestre, par séance de 4 heures au minimum par semaine, les membres du groupe ont à mener des actions d'une semaine à l'autre.

La problématique du projet est clairement définie, mais le choix des solutions techniques pour y répondre est prioritairement laissé à l'initiative des étu-

dants, notamment au niveau BAC+5. Le projet débute donc par une étude bibliographique, qui incite les étudiants à effectuer des recherches de documentations techniques, de normes, etc. tant sur internet, que dans les bibliothèques universitaires. Des pistes leur sont aussi proposées sous forme d'articles scientifiques ou de rapports techniques, leur permettant de se familiariser avec ce type de documents.

Le projet se poursuit par des développements combinant la découverte d'algorithmes de traitement du signal, leur simulation et leur implémentation. La réalisation technique peut en outre prendre des formes variées allant d'une étude de faisabilité à la conception et à la réalisation d'un logiciel, sous Matlab, C ou Java.

Une attention particulière est portée aux moyens de présenter et de valoriser le travail fourni. Cette compétence s'avère essentielle dans le métier d'ingénieur, mais ne peut pas ou difficilement être développée dans un contexte classique d'enseignement. Les étudiants sont donc amenés à présenter leurs résultats lors de réunions d'avancement régulières. Ils sont par ailleurs encouragés à créer un site internet pour lequel ils sont laissés libres du contenu et de l'organisation de la page web avec pour fil directeur la mise en valeur des résultats obtenus. Dans ce cadre, ils peuvent interagir avec l'ingénieur gérant le réseau informatique de l'établissement ou du laboratoire d'accueil des encadrants. Cette partie du projet offre une grande autonomie aux étudiants qui doivent mener une double réflexion sur le choix des résultats présentés et les moyens de les mettre en évidence. Ils se sentent donc particulièrement investis. En outre, ces réalisations rendent leurs compétences plus lisibles pour de futurs employeurs.

Enfin, le projet donne lieu à un rapport en anglais de 4 à 8 pages évalué par un enseignant du département de langues et un rapporteur technique. De plus,

le travail du groupe est valorisé lors d'une soutenance en anglais de vingt minutes en fin de semestre dans un amphithéâtre où tous les étudiants et participants aux projets sont conviés.

Cette forme d'enseignement vise ainsi à favoriser la prise d'initiatives et l'autonomie ; les étudiants qui ont souvent l'habitude de travailler par binôme opèrent dans un mode différent, en interagissant avec plusieurs camarades. Ils doivent donc aussi intégrer les notions intrinsèques de gestion de projet.

Les sujets proposés par l'équipe pédagogique permettent de couvrir une large gamme d'applications (de l'audio aux systèmes de communications mobiles, de l'image à la vidéo, du biomédical au traitement radar, etc.) et peuvent reposer sur des problématiques industrielles. A titre d'exemple, citons :

- Extraction de « signatures » d'un signal de parole pour la reconnaissance de la parole. (BAC +4)
- Introduction au tatouage d'un signal audio (BAC+5)
- Modélisation et simulation d'un canal de propagation de Rayleigh. (BAC +5)
- Mise en œuvre de récepteurs pour les systèmes de communications mobiles pour les systèmes DS-CDMA. (BAC +5)
- Suivi d'ondulation par analyse d'image sur matériaux composites stratifiés. (BAC +5)
- Recalage d'images et mosaïques. (BAC +4)
- Lecture automatique de code-barres dans des séquences vidéo. (BAC +4)
- Suivi et trajectographie d'objets en mouvement, avec application à la circulation routière. (BAC +5)
- Descripteurs pour le suivi d'objets dans des séquences vidéo. (BAC +5)

Remarque 1 : ce type de projet a aussi été mis en place à l'université Bordeaux en Master 2 EEA AP-SI option signal et image. Pour les étudiants qui s'orientent vers une dominante recherche, le sujet du projet est défini pour initier l'étudiant à la recherche. L'accent est alors mis sur l'étude bibliographique.

Remarque 2 : à l'ENSEIRB en filière Télécommunication, ces projets s'inscrivent dans une programmation de projets en traitement du signal sur les trois années de formation. Ainsi, à la fin de la première année, à l'issue des enseignements de base en signal continu et traitement numérique du signal, les étudiants mènent un projet découverte où ils manipulent des signaux biomédicaux ou des signaux de parole qu'ils ont eux-mêmes enregistrés, en général sous Matlab. Puis, durant le premier semestre de la deuxième année, l'ensemble de la promotion travaille sur deux projets : l'un en signal (rehaussement de la parole par soustraction spectrale par exemple qui permet de faire la synthèse entre fenêtrage, stationnarité et ergodicité d'un processus aléatoire et transformée de Fourier rapide), l'autre en traitement d'image (traitement d'image mosaïque, reconnaissance de visages). Dans ce cas, une action spécifique est effectuée durant chaque séance de travail. Les étudiants, organisés par binôme, ne traitent pas nécessairement l'ensemble des étapes du projet.

Enfin, les projets personnalisés débutent le semestre suivant.



Figure 1 : groupe d'étudiants en projet

La suite de cet article se décompose comme suit. Nous détaillons trois applications : la première porte sur le traitement de la parole, la deuxième sur le traitement d'image et la troisième sur le traitement vidéo. A noter que la seconde application met en valeur l'interaction avec le secteur industriel.

1.1 Traitement du signal

Manipuler un signal de parole ou un signal audio est en général une façon ludique de présenter aux étudiants les concepts de traitement du signal 1D, notamment les propriétés fréquentielles, les notions de filtrage ainsi que les problèmes de modélisation paramétrique et d'estimation de paramètres. Dans ce cas, les étudiants obtiennent rapidement des résultats qu'ils peuvent analyser et critiquer, sans qu'ils aient besoin d'une grande expertise a priori.

L'objectif du projet proposé cette année au niveau BAC+4 est plus particulièrement d'étudier les propriétés psychoacoustiques de l'oreille dans un premier temps, puis de les mettre à profit pour se familiariser aux traitements de signaux audio. L'accent est ainsi mis sur les problèmes de masquage fréquentiel qui sont à la base par exemple de la norme MPEG, de techniques de tatouage¹ audio ou de traitements utilisés pour améliorer le rehaussement de la parole. Ce phénomène apparaît lorsque deux sons sont émis simultanément à des fréquences voisines. Il est représenté sur la figure 2 où sont tracés les seuils d'audition lorsque le son masquant est à bande étroite de fréquence centrale 1kHz et de puissance s'échelonnant entre 20 et 100 dB [3].

On peut distinguer deux cas de masquage fréquentiel simultané. Un signal à bande étroite peut être masqué par un autre signal à bande plus large, ou l'inverse.

¹ Le tatouage de signal appelé aussi watermarking se fait sur tous les supports (audio, textes, images et vidéos). Le message codé inséré qui représente par exemple une information de propriété doit être robuste à certains types de traitement.

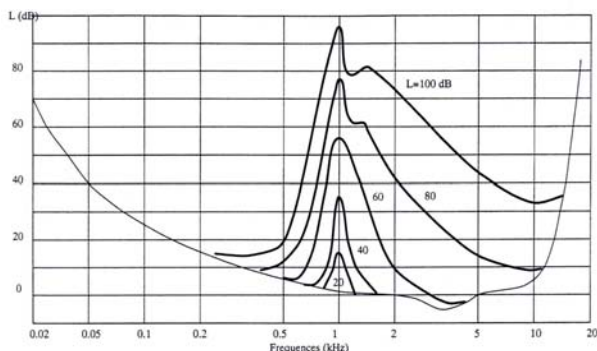


Figure 2 : courbe de masquage fréquentiel résultant d'une tonale produite à 1kHz

Différentes méthodes d'estimation de la courbe de masquage fréquentiel telles que la méthode de Johnston [1] et le modèle psychoacoustique n°1 de MPEG 1 [2] sont donc étudiées par les étudiants et développées sous Matlab. Ce travail est en outre complété par une pré-étude d'un banc filtres pseudo-QMF.

Parallèlement, un autre étudiant a choisi d'aborder des phénomènes complémentaires qui illustrent les propriétés auditives, notamment les illusions auditives et la mélodie du silence en particulier. Ainsi, lorsque plusieurs notes sont jouées simultanément, la mélodie est créée en « arrétant » un court instant l'une des notes du morceau. Celui-ci est donc joué à l'aide des « silences », ce qui peut sembler paradoxal. L'étudiant a alors mis en place une interface originale pour faire appréhender ce phénomène à l'utilisateur : il clique lui-même sur les touches d'un piano virtuel pour créer sa propre mélodie du silence.

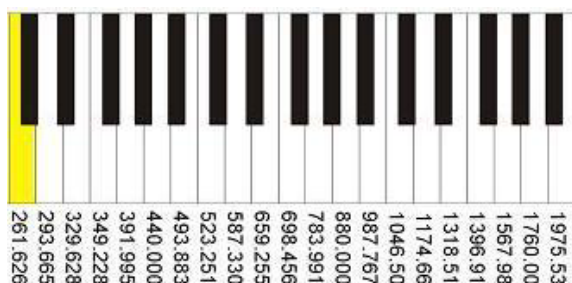


Figure 3 : interface du piano virtuel pour la mélodie du silence

1.2 Traitement d'image

De par nos très forts partenariats, nous disposons d'un certain nombre de thématiques industrielles autour desquelles il est possible de construire des projets accessibles aux étudiants qui sont en phase d'apprentissage des méthodologies de traitement d'images (niveau BAC+4). Il s'agit alors de réaliser des applications restreintes, mais pour lesquelles les étudiants peuvent proposer une solution complète. L'application que nous présentons concerne la mesure de la qualité d'un marquage de type DataMatrix. Ce marquage est utilisé au sein de la société SNECMA (groupe SAFRAN) et est directement apposé sur les composants permettant une complète traçabilité des processus de fabrication (voir Figure 4).

La lecture du code par caméra suppose que le codage soit de bonne qualité. L'objectif du projet est de proposer des mesures automatiques de la qualité du marquage (mesure de distorsion angulaire, mesure de dispersion des rayons des marques, mesure du décentrage des marques par rapport à une grille régulière).

Cette mesure réclame de la part des étudiants la mise en place d'une procédure de filtrage de l'image initiale pour s'affranchir du bruit d'acquisition, suivie d'une procédure de segmentation des marques. Après une étude bibliographique, les étudiants ont opté pour une solution consistant à assimiler les marques à des cercles puis à en rechercher le centre par maximisation d'une mesure de corrélation avec un masque circulaire.

Une fois les centres des marques extraits, le calage d'une grille permet la mesure des critères de distorsion désirés (voir Figure 5).

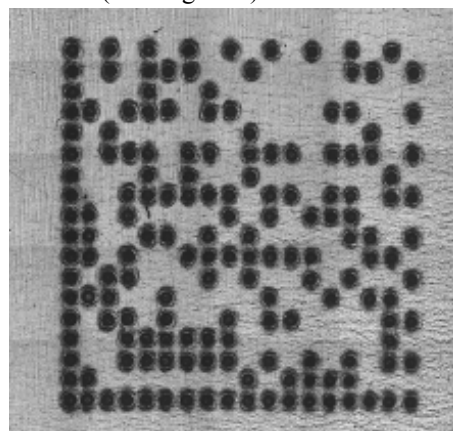


Figure 4 : image de Data Matrix

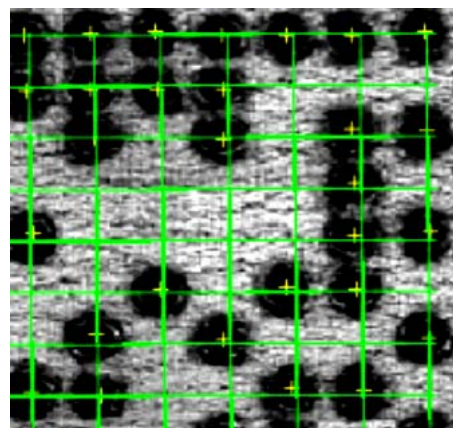


Figure 5 : mise en évidence de distorsions

L'ensemble du processus automatique de mesure de la qualité du marquage a fait l'objet d'une programmation sous Matlab. Cette première version est destinée à être validée avant implantation éventuelle chez notre partenaire, par exemple dans le cadre d'un stage de fin d'étude d'un de nos étudiants.

1.3 Traitement vidéo

Le projet de télécommande virtuelle s'inscrit dans un cadre d'interaction homme-machine, dans laquelle on tire parti des capacités de traitement vidéo de la machine pour proposer de nouvelles façons d'interagir. L'objectif est de développer un système complet

permettant de piloter un certain nombre d'actions pré-définies par l'intermédiaire des mouvements de la main capturés par une caméra [4]. Le système devra détecter la main dans la vidéo, détecter le début et la fin d'un mouvement, suivre la main au cours de ce mouvement, reconnaître le geste accompli et effectuer l'action associée.



Figure 5 : exemple d'image de test et segmentation couleur de la peau associée

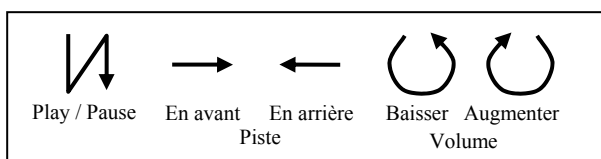


Figure 6 : actions associées aux trajectoires pour la commande d'un lecteur multimédia

Plusieurs aspects du traitement des images et de la vidéo sont abordés ici : la segmentation d'images pour l'extraction des contours de la main [5], le suivi d'objet pour la génération des trajectoires des mains, la reconnaissance de la forme de la main et de la forme de la trajectoire. L'intérêt de cette diversité est de permettre à chaque étudiant de focaliser son travail personnel sur un point particulier du projet, tout en interagissant avec le reste du groupe pour la définition de l'architecture générale du système, et l'intégration des différentes briques au sein d'une application cohérente.

Les notions de traitement d'image vues en cours sont complétées par une étude bibliographique spécifique au sujet proposé, à l'issue de laquelle chaque étudiant choisit une technique adaptée pour le module qu'il gère.

La mise en œuvre de ces nouvelles techniques passe d'abord par un prototype programmé sous Matlab, qui aide à la fois à la validation de la compréhension des algorithmes, à la détermination des paramètres à utiliser (notamment pour la segmentation couleur de la peau), et à la construction d'outils auxiliaires pour la conception du système. Par exemple, dans le cas de la reconnaissance de la forme de la main, un outil permet de définir les descripteurs de Fourier des gestes à reconnaître, qui seront enregistrés dans un fichier pour être utilisés par le système final.

Une fois le prototype et les outils définis, le système final est programmé en langage C, en utilisant l'interface de visualisation N'D développée au sein du groupe signal et image du laboratoire IMS. Chaque étudiant est responsable de son module, un étudiant se consacrant plutôt à l'intégration de l'ensemble. La méthodologie de conception et l'organisation étant particulièrement sensibles pour mener un tel projet à son terme avec succès, ces points sont discutés réguliè-

ment, afin d'aider à la mise en place pratique d'une gestion du projet efficace par les étudiants.

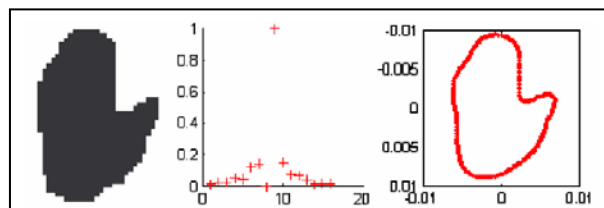


Figure 7 : descripteurs de Fourier pour la reconnaissance de la forme de la main

2 CONCLUSION

Du point de vue des enseignants, ce type de projet s'avère enrichissant à plusieurs titres. Tout d'abord, ces projets sont générateurs d'idées de sujets de TPs en Signal et Image pour les prochaines années. Ils permettent aussi à l'équipe pédagogique d'identifier les faiblesses ou points techniques qui posent des difficultés aux étudiants. Ils constituent donc un « retour » supplémentaire qui permet aux enseignants d'adapter ou de modifier certaines parties des cours classiques, pour mettre en valeur tel ou tel autre concept.

Pour l'étudiant, se focaliser sur un thème particulier et valoriser son travail par une réalisation technique peut s'avérer utile pour préparer un stage et enrichir son CV d'une expérience significative complémentaire.

3 BIBLIOGRAPHIE

- [1] J. D. Johnston, "Transform Coding of Audio Signals Using Perceptual Noise Criteria," IEEE Journal on Selected Areas in Communications, vol. n°6, n°2, pp. 314-323, February 1988.
- [2] N. Moreau : Techniques de compression des signaux. Masson. CENT, Collection Technique et scientifique des Télécommunications, 1994.
- [3] E. Zwicker and H. Fastl, Psychoacoustics, Facts and model, Second updated Edition, Springer Verlag, 1999.
- [4] W. T. Freeman and C.D. Weissman, "Television control by hand gestures", IEEE Intl. Workshop on Automatic Face and Gesture Recognition, Zurich, June, 1995.
- [5] M. J. Jones and J. M. Rehg, "Statistical Color Models with Application to Skin Detection", International Journal of Computer Vision, vol. 46, n°1, pp. 81-96, 2002.