

Technologies multimedias : compression d'images à l'heure des mobiles par Christine Guillemot et Claude Labit. Tatouage d'images, entretien avec Teddy Furon, propos recueillis par Dominique Chouchan.

Christine Guillemot, Claude Labit

► **To cite this version:**

Christine Guillemot, Claude Labit. Technologies multimedias : compression d'images à l'heure des mobiles par Christine Guillemot et Claude Labit. Tatouage d'images, entretien avec Teddy Furon, propos recueillis par Dominique Chouchan.. Les Cahiers de l'INRIA - La Recherche, INRIA, 2008, Les prix Nobel. <inria-00536703>

HAL Id: inria-00536703

<https://hal.inria.fr/inria-00536703>

Submitted on 16 Nov 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TECHNOLOGIES MULTIMEDIAS

Compression d'images à l'heure des mobiles

La vidéo sur téléphone portable devient monnaie courante. Mais la transmission d'images de qualité professionnelle renouvelle profondément la problématique de leur compression et exige des développements mathématiques de plus en plus sophistiqués.

Par définition, les technologies multimédias gèrent des textes, des sons et des images dont la numérisation, notamment celle des images

de compression, Or à l'explosion du nombre d'informations produites et transmises s'ajoute désormais l'extrême hétérogénéité des technologies de transmission et de restitution. Comment les gérer de manière transparente pour l'utilisateur ? Tel est l'un des défis majeurs que nous devons relever aujourd'hui.

Fondamentalement, la compression de données repose sur la théorie de l'information de Shannon, du nom du mathématicien américain Claude Shannon (1916-2001), son fondateur. Cette théorie posait les principes de l'échantillonnage d'un signal, en particulier de la conversion d'un signal analogique (continu) en signal numérique (suite de données numé-

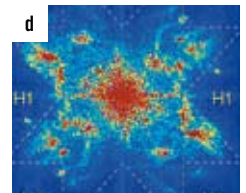
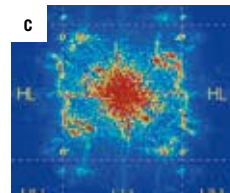
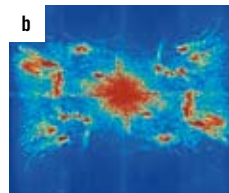


fig. 1 : La compression d'une image (a) passe par sa transformation en signal fréquentiel. Sur les clichés (b), (c) et (d) : les résultats par transformée de Fourier, par décomposition sur une base d'ondelettes classiques et par décomposition sur une base d'ondelettes dites de 2^e génération. On voit que la répartition des coefficients dans toutes les directions de l'espace (en rouge) est nettement plus représentative dans la troisième image.

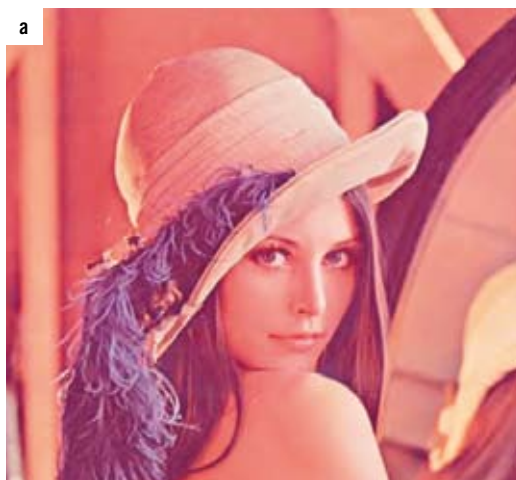
(fixes et animées), se traduit par d'énormes masses de données. Leur archivage et leur transmission *via* un réseau nécessitent donc la mise en œuvre d'algorithmes dits de compression. Ces derniers visent non seulement à réduire le volume des données, sans dégrader l'information, mais aussi à minimiser l'impact des perturbations du réseau lors de la transmis-

sion), et surtout établissait les conditions à respecter pour que l'information ne soit pas trop dégradée. Les procédés de compression de données en découlent directement, les premières réalisations en la matière datant des années 1980. Depuis, les recherches ont conduit à un certain nombre de standards, sous l'impulsion des instances de normalisation*⁽¹⁾.

On aurait pu s'en tenir là si deux éléments nouveaux n'étaient venus bousculer les acquis.

* normes JPEG/MPEG (images fixes/animées) de l'Organisation internationale de normalisation (ISO) et normes telles que H.264 (vidéos) de l'Union internationale des télécommunications (UIT). Lire par exemple: R. Ansari, C. Guillemot and N. Memom, "JPEG and JPEG-2000", *The essential guide to image processing* (Ed. by A. Bovik), Elsevier Inc., 2008

Le premier concerne la diversification des réseaux de transmission des données multimédias, avec l'entrée en scène des réseaux mobiles cellulaires de seconde génération, comme le GSM*, puis de troisième génération, comme l'UMTS*, mais aussi des réseaux informatiques sans fils (Wi-Fi*, Wimax*). Le second élément tient à la diversification parallèle des supports de restitution, notamment des images, avec l'émergence de la vidéo sur téléphone mobile, de la télévision haute définition... L'heure est donc aux réseaux dits multi-technologies ou de quatrième génération.



Une telle infrastructure peut se représenter schématiquement par un ensemble de réseaux d'accès (mobiles ou pas, faible ou haut débit...) à des réseaux dits « cœurs », qui assurent l'interconnexion avec les premiers. Les réseaux cœurs utilisent la technologie par commutation de paquets: ils transmettent des paquets de données (datagrammes), en choisissant à chaque instant et pour chaque paquet les routeurs disponibles, le réassemblage étant réalisé à l'arrivée chez le destinataire. Il faut donc à la fois gérer les différentes caractéristiques de chaque réseau, notamment les bandes passantes (volume de données transmises par seconde), mais aussi la perte éventuelle de paquets... Un seul paquet perdu peut avoir des effets désastreux sur la qualité de service, c'est-à-dire sur les données reçues.

Ce contexte impose de mettre au point des algorithmes qui assurent, outre une forte compression des données, une bonne résistance aux perturbations: on parle de robustesse. Le cas des signaux vidéos, du fait des volumes de données concernés, est paradigmatique. Pour les restituer correctement, il faut au minimum transmettre les contours et la texture de chaque image. Tout l'art consiste à trouver un codage à la fois le plus économe possible en

mémoire et le plus pertinent au plan visuel. C'est pour satisfaire à ces contraintes que nous mettons au point de nouveaux outils mathématiques qui s'appuient à la fois sur la théorie de l'information et sur les techniques de traitement de signal.

En première approximation, la compression peut être imaginée comme une sorte de filtrage: il s'agit de préserver ce qui est significatif. Dans le domaine numérique, une image bidimensionnelle est composée de pixels*. Les algorithmes de compression reposent sur des méthodes d'analyse en fréquence. Celles-ci

consistent à convertir le signal originel (luminance* et chrominance* en chaque pixel) en signal fréquentiel en le décomposant sur une base de fonctions mathématiques: une haute fréquence correspond à une variation rapide de l'image (de couleur, de texture...), une basse fréquence à une zone uniforme. Le signal est alors représenté par un ensemble de coefficients: une zone uniforme sera caractérisée par un coefficient nul aux hautes fréquences tandis que pour une variation rapide de contraste, ce sont les coefficients associés aux basses fréquences qui seront proches de zéro. Comprimer revient à réduire le nombre de coefficients en minimisant la perte d'information.

Depuis une dizaine d'années, les bases de fonctions d'ondelettes se substituent aux bases de fonctions cosinoïdales utilisées auparavant: leur particularité, et leur avantage, sont de s'adapter aux singularités du signal, c'est-à-dire aux zones significatives (contours, contrastes...). La décomposition d'un signal sur une telle base se traduit par un nombre élevé de coefficients dans les régions comportant des contours (plus d'information dans ces régions) et peu de coefficients dans les zones uniformes. Mais le débit nécessaire pour transmettre des images ainsi codées reste trop important.

* **GSM**: Global System for Mobile communications.

UMTS: Universal Mobile telecommunications System.

Wi-Fi: Wireless Fidelity.

Wimax: World Interoperability for Microwave Access

* Un **pixel** est l'unité de surface d'une image numérique

* La **luminance** correspond aux niveaux de gris d'un pixel.

* La **chrominance**, qui comporte deux composantes de l'information couleur, correspond à la couleur du pixel.

* Les **codes correcteurs d'erreurs** sont des techniques de codage visant à corriger les erreurs de transmission d'une information.

Nous travaillons donc aujourd'hui sur de nouvelles bases de fonctions qui rendent compte de l'information dans toutes les directions de l'espace (fonctions anisotropes) avec un nombre de coefficients faible⁽²⁾ (fig. 1).

Il faut en outre que ces fonctions soient adaptées à chaque situation. S'il s'agit de stocker des images, on cherche une représentation aussi parcimonieuse que possible (minimisant le nombre de coefficients). Si l'on veut de surcroît les transmettre *via* un réseau dont la qualité de service n'est pas garantie, il faut anticiper ses perturbations, autrement dit

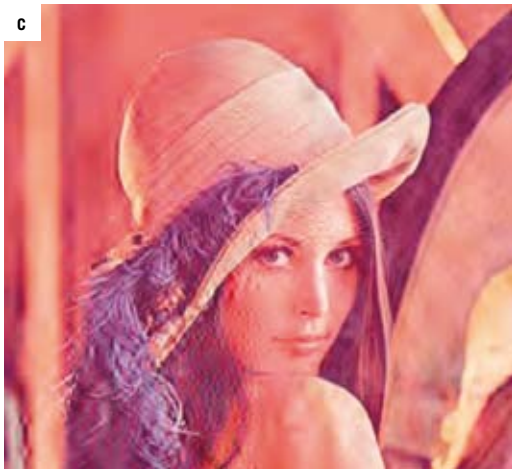


fig. 2 : Le bruit de transmission peut avoir un effet catastrophique sur la qualité du signal décodé chez l'utilisateur. Sur le cliché (a) : une image codée en JPEG-2000 sans erreur de transmission (avec un débit de 0,5 bpp, ou bit par pixel). Sur le cliché (b) : utilisation d'un décodage classique après erreur de transmission. Sur le cliché (c) : décodage robuste après le même type d'erreur de transmission.

développer des algorithmes propres à corriger les erreurs, ces derniers exploitant les propriétés des bases de fonctions d'ondelettes (qui jouent dans ce cas le rôle de codes correcteurs d'erreurs*). Cette étape franchie, un entier est associé à chaque coefficient réel, la transmission d'entiers étant plus efficace que celle de réels. Ces entiers doivent être représentés par des séquences d'éléments binaires aussi courtes que possible, toujours dans le but de compresser au maximum l'information, ce qui suppose d'éliminer toute redondance.

Les coefficients quantifiés (ou parfois les séquences de coefficients quantifiés) sont donc représentés par des suites de bits (appelées mots de code) de longueur variable selon leur probabilité d'apparition. Toutefois, l'inconvénient de tels codes, déjà largement utilisés en compression, est que l'inversion d'un bit (un «0» converti en «1» ou réciproquement) en présence de bruit de transmission peut avoir

un effet catastrophique sur la qualité du signal décodé chez l'utilisateur (fig. 2). Dans le cadre du réseau d'excellence européen Newcom++ (7^e PCRD), mis en place au début de l'année 2008, nous développons donc de nouveaux codes, dits *joint source-channel* : l'idée est de maintenir une redondance contrôlée dans la représentation binaire, cette information redondante servant à corriger le bruit ou les erreurs introduites par le canal de transmission⁽³⁾.

Au final, l'objectif est par exemple de mettre en œuvre des services tels que la vidéo à la demande. Cela suppose de compresser les données une fois pour toutes et de pouvoir les transmettre à différents débits et les décoder à différentes résolutions spatiales et temporelles, selon les caractéristiques du récepteur. Mais comment garantir une telle adaptation de manière dynamique, et surtout totalement transparente pour l'utilisateur ? Nous sommes en train de mettre au point des outils algorithmiques complémentaires de compression dits «scalables» (ou «échelonnables»). Une première solution normalisée (le standard MPEG-SVC, pour Scalable Video Coding) vient d'être finalisée par le groupe *ad hoc* de l'ISO et de l'UIT mais elle ne répond pas complètement aux besoins applicatifs.

Enfin, la restitution d'images tridimensionnelles (3D) a pris un véritable essor, tant dans le domaine du cinéma numérique que de la télévision, grâce à la disponibilité de systèmes d'acquisition et d'écrans stéréoscopiques à des coûts modérés. Pour compresser ces images, il s'agit d'exploiter les redondances entre vues capturées par plusieurs caméras, d'en extraire des informations sur la profondeur des objets présents dans les scènes et de modéliser celles-ci en trois dimensions⁽⁴⁾. Les critères de robustesse et de scalabilité doivent alors garantir une bonne restitution du relief et une exploitation optimale des informations 3D pour générer de nouvelles vues, permettre de naviguer dans les images, voire pour mixer des objets réels et virtuels (réalité augmentée). C'est un immense chantier de recherche qui s'ouvre pour étendre les concepts et algorithmes de compression évoqués précédemment.

Christine Guillemot, directrice de recherche INRIA (Rennes) et responsable de l'équipe projet TEMICS, est membre de plusieurs comités scientifiques internationaux et éditrice associée de la revue internationale de traitement du signal *IEEE Trans. on signal Processing*.

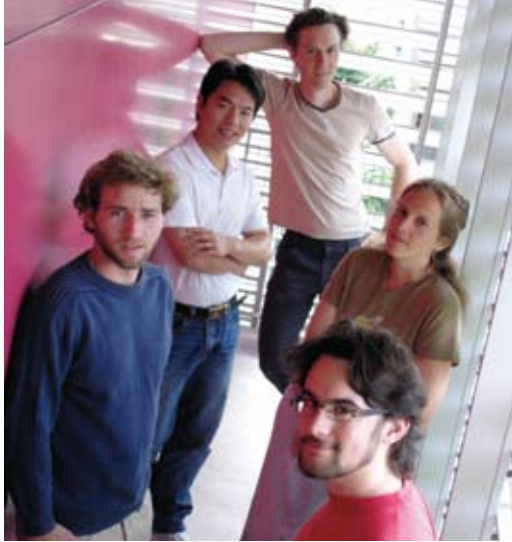
Claude Labit, directeur de recherche INRIA et cocréateur de l'équipe projet TEMICS, a dirigé le centre INRIA de Rennes de 1999 à 2007 et préside aujourd'hui la technologie de Rennes-Atalante.

(1) M. Barlaud et C. Labit (Ed.), «Compression et codage des images et vidéos». *Traité IC2, série Traitement du signal et de l'image*, Ed. Hermès, 2002.

(2) V. Chappelier et C. Guillemot, «Oriented wavelet transform for image compression and de-noising», *IEEE Trans. on Image Processing*, vol. 15 (10), p. 2892, 2006.

(3) C. Guillemot et P. Siohan, «Joint source-channel decoding with soft information : a survey», *Eurasip journal on applied signal processing*, vol. 6, p. 906, 2005.

(4) O. Scheer, P. Kauff et T. Sikora (Ed.), «3D video-communication : Algorithms, concepts and real-time systems», Ed. John Wiley, 2005.



© CNRS-N.TIGET

Alors que se multiplient les techniques de diffusion des images et des sons, celles pour anticiper ou combattre le piratage ne cessent de se complexifier.

De haut en bas : **Teddy Furon**, chargé de recherche INRIA (équipe projet TEMICS) et spécialiste des techniques de protection des contenus multimédias, avec ses collègues **Fuchun Xie** et **Ana Charpentier** (doctorants), **Mathieu Desoubeaux** (ingénieur) et **Philippe Roudot** (stagiaire).

Entretien avec Teddy Furon

Tatouage d'images

Pourquoi tatouer des images ?

Teddy Furon : Le tatouage consiste à cacher des informations dans un contenu (visuel, sonore...). Tel un filigrane, il est invisible (ou inaudible). Ces informations sont enfouies dans le contenu au point d'en devenir indissociables. L'une des applications est de protéger les droits de propriété intellectuelle autour d'une œuvre en insérant le nom de l'auteur ou bien ses conditions de diffusion. Un autre moyen d'assurer cette protection est d'inscrire un code identifiant le ou les acheteur(s) de l'œuvre, afin de suivre à la trace sa rediffusion illégale éventuelle. Autre application possible : cacher un numéro type ISBN* dans des créations sonores afin d'évaluer leur fréquence de diffusion. On pourrait même imaginer des utilisations cliniques, à l'heure où toutes les images médicales sont sous forme numérique, le but étant alors d'éviter toute interversion d'images de patients.

Quels sont les critères d'évaluation d'un « bon » tatouage ?

T. F. : Le tatouage est bien sûr imperceptible. Il est aussi robuste : on retrouve les informations cachées même si l'image ou le son sont dégradés. D'une manière générale, le choix d'un tatouage est affaire de compromis. Par exemple, c'est la répétition du tatouage qui crée la robustesse : plus le message à cacher est court, plus il peut être répété dans l'image et

D'UNE FRONTIÈRE À L'AUTRE

meilleure est la robustesse. Il faut donc faire un compromis entre taille du message et robustesse. Autre cas de figure : le cinéma numérique. Les producteurs souhaitent insérer un tatouage à chaque projection (heure et salle) de manière à localiser un éventuel « pirate » qui aurait « camcordé* » la projection du film. Dans ce cas, le compromis est entre complexité du message et contraintes temporelles : il faut tatouer en temps réel des images de très bonne qualité et très grandes.

Pourquoi, aujourd'hui, sécurise-t-on les tatouages ?

T. F. : La prise en compte de la dimension sécurité du tatouage est très récente. Jusqu'à il y a peu, on pensait que le tatouage offrait par lui-même une garantie de sécurité, ce qui est loin

d'être le cas. D'où l'idée de le chiffrer par une clé secrète. Mais comment savoir si un petit malin ne va pas aussi casser la clé ? Autrement dit, après observation de combien d'images la retrouvera-t-il ? Nos recherches portent aujourd'hui sur l'augmentation du niveau de sécurité. Autre question : si un identifiant (associé à chaque client) est caché dans une production visuelle, et si quelques clients sont malhonnêtes (ils fusionnent leurs contenus de manière à brouiller les informations cachées puis diffusent l'œuvre), comment retrouver la copie d'origine donc les responsables ? C'est l'un de mes thèmes de recherche actuels. Un contrôle *a posteriori* donc, mais dissuasif. Il s'agit alors de discriminer des signaux extrêmement ressemblants tout en minimisant la probabilité d'accuser un innocent.

La compression, en dégradant le signal, ne vous complique-t-elle pas la tâche ?

T. F. : Il est très difficile d'anticiper l'impact de la compression sur le tatouage. Le raisonnement consiste à dire que la compression n'est pas l'attaque la plus dangereuse, au regard d'attaques, dites optimales, beaucoup plus néfastes pour le tatouage. Donc en se préparant au pire, on espère que la compression ne sera pas un problème. C'est ce que nous avons fait dans notre équipe en nous appuyant notamment sur la théorie des jeux.

Propos recueillis par Dominique Chouchan

* L'*International Standard Book Number*, ou ISBN, désigne un numéro destiné à identifier tout livre publié dans le monde.

* le terme « camcorder » est un anglicisme issu du néologisme anglais *camcorder* construit à partir des mots *camera* et *recorder*. Un *camcorder* est un caméscope.