Techniques for improving the performance of distributed video coding

Abdalbassir Abou-Elailah Abou El Ailah

▶ **To cite this version:**

Abdalbassir Abou-Elailah Abou El Ailah. Techniques for improving the performance of distributed video coding. Other. Télécom ParisTech, 2012. English. NNT : 2012ENST0075 . tel-00794685v2

HAL Id: tel-00794685

https://pastel.hal.science/tel-00794685v2

Submitted on 14 Jun 2013

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.

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

TELECOM ParisTech

2012-ENST-075

EDITE - ED 130

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

TELECOM ParisTech

Spécialité « Signal et Images »

présentée et soutenue publiquement par

Abdalbassir ABOU-ELAILAH

le 14 décembre 2012

Techniques d'amélioration des performances de compression dans le cadre du codage vidéo distribué

Directeur de thèse : **Frederic DUFAUX**
Co-encadrement de la thèse : **Marco CAGNAZZO**

Jury

M. Touradj EBRAHIMI, Professeur, EPFL	Président
M. Fernando PEREIRA, Professeur, Instituto Superior Técnico	Rapporteurs
Mme Lina KARAM, Professeur, Arizona State University	
M. Eric NASSOR, Manager, Canon Research Centre France	Examinateur
M. Frederic DUFAUX, Directeur de recherche CNRS, Télécom ParisTech	Directeurs de Thèse
M. Marco CAGNAZZO, Maitre de conférence, Télécom ParisTech	
Mme Beatrice PESQUET-POPESCU, Professeur, Télécom ParisTech	
Mme Joumana FARAH, Professeur, Université Saint-Esprit De Kaslik	

TELECOM ParisTech
école de l'Institut Télécom - membre de ParisTech

THÈSE

Remerciements

Je tiens à adresser mes chaleureux remerciements à mon directeur de thèse Dr. Frédéric DUFAUX, pour avoir dirigé d'une façon continue mes activités de recherche. J'ai profondément apprécié les discussions constructives avec lui ainsi que la confiance qu'il m'a accordée pendant ce travail de thèse. J'ai également apprécié ses relectures, remarques, commentaires et corrections de mes travaux de recherche.

Je remercie très sincèrement mon co-directeur de thèse Dr. Marco CAGNAZZO pour la proposition de mon sujet de thèse, pour l'encadrement de mes activités de recherche et pour la riche compréhension qu'il m'a apportée du codage vidéo distribué.

Je tiens également à remercier Prof. Béatrice PESQUET-POPESQU, chef du groupe MultiMedia à Telecom Paristech, pour son accueil chaleureux dans son groupe de recherche, pour l'encadrement professionnel de mes activités de recherche et pour ses relectures et remarques constructives de mes travaux.

Je souhaite remercier très chaleureusement Prof. Joumana FARAH, pour avoir encadré mon travail de recherches à distance, pour sa confiance, ses encouragements et ses conseils techniques et pratiques. J'ai particulièrement apprécié les discussions avec elle, les relectures et les corrections précieuses de mes articles et de mon manuscrit.

Mes remerciements s'adressent aux rapporteurs : Prof. Lina KARAM et Prof. Fernando PEREIRA et aux examinateurs : Prof. Touradj EBRAHIMI (président du jury) et Dr. Eric NASSOR qui avoir accepté d'évaluer et de rapporter mon travail de recherches, et pour leur participation au Jury de soutenance.

Je remercie aussi Prof. Michel KIEFFER et Dr. Joel JUNG pour avoir effectué l'évaluation de mi-parcours de mon travail de thèse.

Je tiens à remercier Eng. Julien LE TANOUR, pour son aide et sa collaboration dans la compréhension et l'application des courbes élastiques.

Je voudrais également remercier le personnel technique et administratif de Telecom Paristech, pour leur support pendant ma thèse.

Ma gratitude s'adresse également à mes collègues durant ma thèse à Telecom Paricteh : Thomas, Giovanni, Claudio, Rafael, Mounir, Irina, Manel, Elie, Hamlet, Giuseppe, Marc, Paul et Alper qui m'ont aidé à répondre à un certain nombre de questions pratiques, techniques et scientifiques.

Enfin, un grand merci à mes parents, mes frères, mes sœurs et mes amis qui sont

toujours à mes côtés, et qui m'ont continuellement soutenu pendant ces années de travail de thèse.

Abdalbassir ABOU-ELAILAH

Résumé

Le codage vidéo distribué (DVC) est une technique récemment proposée dans le cadre du codage et de la transmission des séquences vidéo, et qui convient surtout à une nouvelle classe d'applications telles que la surveillance vidéo sans fil, les réseaux de capteurs multimédia, les caméras PC sans fil, les téléphones mobiles et les appareils-photos numériques. Ces applications nécessitent en effet un encodeur de faible complexité, avec la possibilité d'un décodeur de complexité élevée. DVC présente plusieurs avantages : d'abord, la complexité peut être distribuée entre l'encodeur et le décodeur. Deuxièmement, le DVC est robuste aux erreurs, car un codeur de canal y est incorporé. En DVC, une information adjacente (Side Information ou SI) est estimée au décodeur en se basant sur les trames décodées disponibles, et utilisée pour le décodage et la reconstruction des autres trames.

Dans cette thèse, nous proposons de nouvelles techniques qui permettent d'améliorer la qualité de l'information adjacente. Tout d'abord, le raffinement itératif de l'information adjacente est réalisé après le décodage de chaque sous-bande DCT, en utilisant la trame Wyner-Ziv (WZ) partiellement décodée (appelée PDWZF) avec les trames de référence. De plus, dans cet algorithme, une nouvelle approche est proposée qui permet d'adapter la fenêtre de recherche au niveau de mouvement courant entre la trame WZ et les trames de référence, en se basant sur la PDWZF obtenue après le décodage de la première sous-bande DCT. Ensuite, une nouvelle méthode de génération de l'information adjacente est proposée, qui utilise l'estimation des vecteurs de mouvement dans les deux sens et le raffinement Quad-tree. En outre, en vue d'améliorer la qualité des trames WZ décodées pour les grandes tailles de GOP (Group Of Pictures), un algorithme basé sur les trames adjacentes décodées est proposé, qui utilise une zone de recherche adaptative et une taille de bloc variable.

Une autre contribution de cette thèse concerne la fusion des estimations globale et locale. Les paramètres globaux sont calculés au codeur en utilisant l'algorithme SIFT. Ces paramètres globaux sont transmis au décodeur pour y être utilisés dans l'estimation de l'information adjacente globale. Ensuite, de nouvelles approches sont proposées afin de combiner les estimations de mouvement globale et locale. Dans la première approche, la fusion se base sur les différences entre les blocs correspondants. Dans la seconde, la technique SVM (Support Vector Machine) est utilisée pour combiner les deux informations adjacentes. En plus, des algorithmes sont proposés pour améliorer la fusion au cours du

décodage, par l'exploitation de la PDWZF et des coefficients DC décodés. En outre, les objets segmentés des trames de référence sont utilisés dans la combinaison des estimations de mouvement globale et locale, en utilisant les courbes élastiques et la compensation de mouvement basée-objets.

De nombreuses simulations ont été effectuées pour tester les performances des techniques proposés et qui montrent des gains importants par rapport au codeur classique DISCOVER. Par ailleurs, les performances de DVC obtenues en appliquant les algorithmes proposés surpassent celles de H.264/AVC Intra et H.264/AVC No motion pour les séquences testées. En plus, l'écart vis-à-vis de H.264/AVC Inter avec une configuration IB...IB est considérablement réduit.

Abstract

Distributed Video Coding (DVC) is a recently proposed paradigm in video communication, which fits well emerging applications such as wireless video surveillance, multimedia sensor networks, wireless PC cameras, and mobile cameras phones. These applications require a low complexity encoding, while possibly affording a high complexity decoding. DVC presents several advantages: First, the complexity can be distributed between the encoder and the decoder. Second, the DVC is robust to errors, since it uses a channel code. In DVC, a Side Information (SI) is estimated at the decoder, using the available decoded frames, and used for the decoding and reconstruction of other frames.

In this Ph.D thesis, we propose new techniques in order to improve the quality of the SI. First, successive refinement of the SI is performed after each decoded DCT band, using a Partially Decoded WZF (PDWZF), along with the reference frames. Moreover, in this refinement approach an adaptive search area algorithm is also proposed, that allows adapting the search area to the current motion between the WZF and the reference frames, using the PDWZF obtained after decoding the first DCT band. Then, a new scheme for SI generation based on backward, forward motion estimations, and Quad-tree refinement is proposed. Furthermore, in the aim of enhancing the quality of the decoded WZFs for larger GOP sizes, an algorithm based on adjacent decoded frames is investigated, using an adaptive search area and a variable block size.

Another contribution of this thesis concerns a fusion of global and local SI. Global parameters are estimated at the encoder using the Scale-Invariant Feature Transform (SIFT) algorithm. These global parameters are sent to the decoder to estimate the global SI. Then, new methods for combining global and local motion estimations are proposed, to further improve the SI. In the first approach, the differences between the corresponding blocks are used to combine the global and local SI frames. In the second approach, Support Vector Machine (SVM) is used to combine the two SI frames. In addition, algorithms are proposed to refine the fusion during the decoding process by exploiting the PDWZF and the decoded DC coefficients. Furthermore, the foreground objects are used in the combination of the global and local motion estimations, using elastic curves and foreground objects motion compensation.

Extensive experiments have been conducted showing that important gains are obtained by the proposed techniques compared to the classical DISCOVER codec. In addition, the

performance of DVC applying the proposed algorithms outperforms now the performance of H.264/AVC Intra and H.264/AVC No motion for tested sequences. Besides that, the gap with H.264/AVC in an Inter IB...IB configuration is significantly reduced.