



Title	A Study on Multi-view Video Crowdsourcing over Wireless Networks
Author(s)	Nu, Than Than
Citation	大阪大学, 2020, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/76649
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

Abstract of Thesis

Name (Than Than Nu)

Title

A Study on Multi-view Video Crowdsourcing over Wireless Networks
(無線ネットワークを介した多視点ビデオクラウドソーシングに関する研究)

Abstract of Thesis

The integration of video streams captured by many mobile video contributors at a crowded event into a multi-view video, that is, crowdsourced multi-view video streaming enables the remote viewers to experience the immersive views of the event as if they were attending the event. The serviceability of crowdsourced multi-view videos streaming is not limited to entertainment purposes. As an example, it can be extended to intelligent transportation system (ITS) in which the videos captured by onboard cameras of different vehicles at one side of an intersection can be shared with the vehicles at the other side to provide the traffic condition at the intersection. Another potential application of crowdsourced multi-view video streaming is three-dimensional (3D) virtual reconstruction of an interesting location, building, or a concert using video frames to point cloud technology. Despite the diverse applicability, the crowdsourcing multi-view videos over wireless network encounters many challenges due to the resource-constrained nature of wireless networks and limited capacity of the consumer-grade mobile devices.

One of the characteristics of crowdsourced multi-view streaming is that many contributors upload the video streams captured at close spatial location simultaneously to the video collector, that is, a server. In this case, a large amount redundant video traffic is uploaded to the serve due to the high correlation among the video streams, which is highly inefficient usage of scarce wireless network resources. In this case, traffic reduction is of paramount importance to realize the efficient of the uploading of crowdsourced video streams. On the other hand, due to the power-hungry nature of video streaming and limited battery supply of mobile devices restrict the practicability of crowdsourced multi-view video streaming. Thus, it is demanding that the traffic-reduced, high-quality and low-power video uploading solutions to realize the efficient crowdsourced multi-view video streaming.

This dissertation proposes three video uploading schemes considering the above-mentioned issues. Firstly, the author proposes a traffic reduction method for multi-view video uploading from crowdsourced video contributors. The proposed scheme uses differential encoding with multiple reference streams by means of packet overhearing. To realize differential encoding across the network of contributors for higher traffic reduction, the proposed scheme combines three techniques: correlation estimation, reference selection, and transmission order determination. First, the scheme utilizes the correlation among the contributors based on the content features of the captured video streams using the information-bound reference (IBR). Second, in the design of the reference selection that determines the dependencies among the contributors, two threshold values are used, determining the number of references for differential encoding at each contributor. Finally, the scheme schedules the transmission order of the contributors to increase the number of differential encoding opportunities within their network.

Second, the author proposes a cluster-based redirect video uploading scheme for high-quality and low-power crowdsourced multi-view video streaming. Considering the drawbacks of conventional digital video transmission consumption, the proposed scheme integrates the four approaches of network clustering, delegate selection, soft video delivery, and four-dimensional discrete cosine transform (4D-DCT) to redirectly upload the captured videos to the AP. Specifically, network clustering and delegate selection leverage the redirect path between the contributors and the AP. Soft video delivery removes power-hungry digital encoding and transmission by directly sending frequency-domain coefficients using multi-dimensional DCT and near-analog modulation. 4D-DCT exploits the content correlations between the contributors to reduce redundant transmissions.

Finally, the author extends the previously proposed differential encoding-based video uploading scheme using ROI-based multi-view video encoding with the goal of further traffic reduction considering the perceptual redundancy of human visual system (HVS). To realize the ROI-based differential multi-view video uploading, firstly, a correlated network of contributors is constructed based on the correlation degrees among them. Then the proposed scheme selects a contributor with the largest average correlation in the network for the task of extraction of ROI regions. The selected contributor uploads its video stream using ROI-based single-view video encoding while the rest of the contributors in the network overhear the transmitted stream and performs ROI-based differential multi-view video encoding with the prior knowledge ROI information.

This dissertation evaluates the performances of the proposed schemes through computer simulations implemented in MATLAB. The standard multi-view video sequences provided by Mitsubishi Electronic Research Laboratory (MERL) are used for evaluations.

According to the simulations results, the proposed traffic reduction scheme achieves a traffic reduction of up to 31% with a quality improvement of up to 2.7 dB as compared to the conventional video uploading scheme in the densely correlated network of contributors. The second proposed scheme which leverages the redirect soft video uploading outperforms the conventional digital-based video delivery in terms of video quality and transmission power consumption by skipping power-hungry digital video encoding. The last proposed scheme, ROI-based differential encoding scheme achieves 57% of traffic reduction with the unnoticeable quality of degradation in the less interested regions by exploiting the perceptual redundancy of human vision system.

論文審査の結果の要旨及び担当者

氏 名 (Than Than Nu)		
論文審査担当者	(職)	氏 名
	主査 教授	渡辺 尚
	副査 教授	村田 正幸
	副査 教授	長谷川 亨
	副査 教授	東野 輝夫
	副査 教授	松岡 茂登

論文審査の結果の要旨

複数の携帯機器によって撮影された映像を利用するクラウドソース型マルチビュービデオストリーミングは、臨場感あふれる種々のアプリケーションに応用可能である。例えば、3次元スポーツ観戦や仮想コンサート等の娯楽、複数の車載ドライブレコーダや路肩カメラの映像を駆使する高度なインテリジェント交通システム、複数アングルからの構造物モニタリング等を挙げることができる。

クラウドソース型マルチビュービデオストリーミングでは、群衆がそれぞれ携帯する端末によって映像を撮影してアップロードする。従って、無線ネットワークを介した複数映像の収集は無線ネットワークおよび携帯端末の制約に起因する多くの課題を抱えている。特に、映像品質の劣化はクラウドソース型マルチビュービデオストリーミングの成否を左右する重要な性能指標である。

クラウドソーシング型マルチビュービデオストリーミング映像の高品質化を達成するためには、1) 無線ネットワークのトラヒックを削減すること、2) 端末の消費電力を削減することが必要となる。無線ネットワークのトラヒック削減のためには、高度な映像符号化技術を用いることとなる。しかしながら、このことは携帯端末の著しい電力消費を強いることとなる。すなわち、トラフィック削減と消費電力削減はトレードオフの関係にあり、両者のバランスを解決する無線映像伝送手法が不可欠である。

本論文は、この課題に対し、無線映像伝送の新たな3手法を提案し、その効果を理論的に論じたものである。第1の手法では、携帯端末間で映像情報を含むパケットをオーバーヒアして得られた他端末の映像情報を用いた差分エンコーディングを実現することによりトラヒック削減を達成する。より大幅なトラヒック削減を図るため、本手法では端末間の類似度を推定するとともに、どの端末から映像情報を得るべきかを選択する2種のアルゴリズムを提案している。

第2の手法は、4次元離散コサイン変換およびニアアナログ変調を組み合わせて映像符号化を必要としない無線映像伝送を実現することで低消費電力化を達成する手法である。具体的には、4次元離散コサイン変換を送信時に利用することで空間・時間・撮影機器間に含まれる冗長な映像情報を削減する。

第3手法は、人間は限られた領域を注視する特性を利用し、各撮影機器に含まれる冗長な映像情報を削減することにより、さらなるトラヒック削減を達成する方法である。この手法では注視領域を考慮して、注視領域内に含まれる映像情報を優先的に符号化し、映像収集の高品質化を達成する。

これらの提案手法をソフトウェアと実映像を利用したシミュレーションにより評価した。その結果から、提案手法は従来手法と比較して低トラヒックを達成することを確認した。また、第2の提案手法は映像符号化に起因する電力消費量を大幅に削減できることを明らかにした。

以上の成果は、次世代映像配信技術の高度化・省エネルギー化を支える根幹技術であり、学術的価値が高くまた社会へ大きく貢献する。よって、博士（情報科学）の学位論文として価値あるものと認める。