

OPTIMIZING PEER GROUPING FOR LIVE FREE VIEWPOINT VIDEO STREAMING

Yuan Yuan^{*}, Bo Hu^{*}, Gene Cheung[#], Vicky Zhao^{*}

^{*} University of Alberta, Canada, [#] National Institute of Informatics, Japan

ABSTRACT

In free viewpoint video, a user can pull texture and depth videos captured from two nearby reference viewpoints to synthesize his chosen intermediate virtual view for observation via depth-image-based rendering (DIBR). For users who are observing the same video at the same time but not necessarily from the same virtual viewpoint, they have incentive to pull the same reference views so that the streaming cost can be shared. On the other hand, in general distortion of a synthesized virtual view increases with its distance to the reference views, and so a user also has incentive to select reference views that tightly “sandwich” his chosen virtual view, minimizing distortion. In a previous work, reference view sharing strategies—ones that optimally trade off shared streaming costs with synthesized view distortions—were investigated for the case when users are first divided into groups, and each user group independently pulls two reference views and shares the resulting streaming cost. In this paper, we generalize the previous notion of user group, so that a user can simultaneously belong to two groups, and each group shares the streaming cost of a single view. We also aim to find a Nash Equilibrium (NE) solution of reference view selection, which is stable and from which no one has incentive to unilaterally deviate. Specifically, we first derive a lemma based on known properties of synthesized view distortion functions. We then design a search algorithm to find a NE solution, leveraging on the derived lemma to reduce search complexity. Experimental results show that the stable NE solution increases the overall cost only slightly when compared to the unstable optimal reference selection that gives the lowest overall cost. Further, a larger network will give a lower average cost for each user, and thus, users tend to join large networks for cooperation.

Index Terms— Free viewpoint video, content sharing, Nash equilibrium

1. INTRODUCTION

In free viewpoint video [1], a user can select any virtual view from which an image of the 3D scene is rendered for observation. Specifically, given a 1D array of cameras with positions $\mathcal{V} = \{1, \dots, V\}$, an image of virtual view u is typically synthesized using texture and depth maps captured from two nearby captured views, v^l and v^r , where $v^l < u < v^r$ and $v^l, v^r \in \mathcal{V}$, via depth-image-based rendering [2]. For users who are observing the same free viewpoint video synchronized in time—e.g., during a live video broadcast of a public event like a piano recital—but not necessarily from the same viewpoint, they have incentive to pull texture and depth video streams from the same reference views, so that the streaming cost can be shared. On the other hand, it has been shown [3, 4] that in general distortion of the synthesized view increases with its distance to the reference views. Thus, a user also has incentive to select video of reference views that tightly “sandwich” his chosen virtual view, in order to minimize visual distortion. This poses an interesting dilemma for

users: how to best select and share video streams of different reference views, so that the streaming cost and the resulting collective synthesized view distortion is optimally traded off?

In a previous work [5], reference view sharing strategies were studied for the case where users are first divided into groups, and then each group independently pulls and shares the streaming cost of two reference views, using which virtual views of the group’s users are synthesized. While the developed algorithms are simple and intuitive, it is easy to see how this type of groupings is sub-optimal. First, it is possible for multiple groups to be independently pulling the same video view, when the cost of this common view can be shared by the union of these groups. Second, members belonging to the same group must share *both* reference views, when it may be more optimal for them to share only one reference view, and separately find appropriate groups to share a different second reference view for view synthesis. By imposing the constraint that each user group selects two reference views only for users in that group, neither of these two cases are possible.

In this paper, we generalize the previous notion of user group, so that a user can simultaneously belong to two groups, and each group shares the streaming cost of a single view. Doing so means a video view is never pulled more than once, and its cost is shared only by those who are using this view as reference for view synthesis. To study a stable user grouping, we exploit tools from game theory [6], and seek a Nash Equilibrium (NE) solution of reference view selection, from which no one has incentive to unilaterally deviate. Specifically, we first derive a lemma based on known properties of synthesized view distortion functions. We then design a search algorithm to find locally optimal groupings, leveraging on the derived lemma to reduce search space, thus reducing computation complexity. Simulation results show that the stable NE solution achieves slightly higher overall cost than the unstable optimal reference selection that gives the lowest overall cost. Furthermore, a larger network has the ability to request more reference views to reduce users’ distortion without much increase of the subscription fees shared by each user. Thus, everyone can reach a lower cost in a larger network.

The outline of the paper is as follows. We first overview related work in Section 2. We then formulate our problem in Section 3. We derive our lemma and corresponding optimization algorithm in Section 4. Finally, experimental results and conclusion are presented in Section 5 and 6, respectively.

2. RELATED WORK

As technologies for compression of texture and depth maps for free viewpoint video become more mature [7, 8], research focus has shifted to the streaming and distribution of this new media type. [9] designed a multiview video compression algorithm in combination with an observer’s head position prediction scheme, so that the likely captured video views to be observed by client in the near future are automatically pulled from server. [10, 11] studied the problem of

how smartly encoded multiview video that facilitates view-switching can be replicated in storage-constrained distributed servers across a network to minimize view-switching delay. [12, 13] investigated how texture and depth videos can be unequally protected to minimize the synthesized view distortion when streaming over a network prone to packet losses. None of these prior streaming work studied the problem of how video streams of different views can be optimally selected and shared among users observing different virtual views, however, which is the focus of this paper.

On the other hand, video sharing for single-view video, mostly for Peer-to-Peer (P2P) video streaming, has been studied extensively in the literature. For example, the work in [14] derived a stochastic fluid model to analytically reveal the characteristics of P2P streaming systems and exposed the key designing features to achieve a satisfactory system performance. The work in [15] studied a real world large-scale P2P streaming system to gain insights for successful deployment of such systems. The work in [16] reviewed different overlay network structures for both P2P live streaming and video-on-demand. However, all these works for single-view streaming cannot be directly applied to the free viewpoint scenario, since how to select and share the reference views to address the tradeoff between the streaming cost and the synthesized view distortion is a key issue for live free viewpoint video distribution, which we study here.

3. PROBLEM FORMULATION

In this section, we first describe the free viewpoint video model we chose for our problem formulation. We then describe properties of the synthesized view distortion and subscription fee sharing.

3.1. Free Viewpoint Video Model

Let $\mathcal{V} = \{1, \dots, V\}$ be a discrete set of *captured views* for V equally spaced cameras in a 1D array. Each camera captures both a texture map (RGB image) and a depth map (per-pixel physical distances between objects in the 3D scene and camera) at the same resolution. The texture map from an intermediate *virtual view* between any two cameras can be synthesized using texture and depth maps of the two camera views (*reference views*) via a depth-image-based rendering (DIBR) technique like 3D warping [2]. DIBR essentially maps texture pixels in the reference views to appropriate pixel locations in a virtual view; such locations are derived from the corresponding depth pixels in the reference views. Disoccluded pixels in the synthesized view—pixel locations that are occluded in the two reference views—can be completed using depth-based inpainting techniques [17, 18]. Because inpainting offers only a best-guess solution, the larger the disoccluded regions are, the lower the synthesized view image quality will be in general.

More specifically, let u be the virtual view that a peer currently requests for observation. We assume u can be written as $u = v + \frac{k}{K}$, $v \in \{1, \dots, V - 1\}$ and $k \in \{0, \dots, K\}$, for some large predetermined constant K . In other words, u belongs to an ordered discrete set of intermediate viewpoints—the set of views between (and including) camera views 1 and V , spaced apart by integer multiples of $1/K$. A discrete distribution function q_u describes the fraction of peers who currently request the virtual view u .

3.2. Synthesized View Distortion

Typically, to construct a virtual view u that is not itself a camera-captured view, DIBR requires left and right reference views v^l and v^r such that $v^l < u < v^r$. Note that v^l and v^r do not have to be the closest captured views to u . The distortion of the synthesized view, $d_u(v^l, v^r)$, varies with the choices of reference views, v^l and v^r . We assume that $d_u(\cdot)$ has the following three properties.

First, further away reference views v^l and v^r to virtual view u induce no smaller distortion, that is,

$$\begin{aligned} d_u(v^l, v_1^r) &\leq d_u(v^l, v_2^r) \text{ if } v_1^r < v_2^r, \text{ and} \\ d_u(v_2^l, v^r) &\geq d_u(v_1^l, v^r) \text{ if } v_2^l < v_1^l. \end{aligned} \quad (1)$$

We call this the *monotonicity in reference view distance* for synthesized view distortion. This is reasonable, since further reference views usually result in more disoccluded pixels, lowering the quality of the synthesized view [3, 4].

Second, given virtual view u and left and right reference views v^l and v^r , define $\tau = \min(|v^l - u|, |v^r - u|)$ as the *minimum reference view distance* between u and v^l, v^r . We assume that a smaller τ induces no larger distortion, i.e.,

$$\begin{aligned} d_{u_1}(v^l, v^r) &\leq d_{u_2}(v^l, v^r) \text{ if } \tau_1 < \tau_2 \\ \text{where } \tau_1 &= \min(|v^l - u_1|, |v^r - u_1|) \\ \text{and } \tau_2 &= \min(|v^l - u_2|, |v^r - u_2|). \end{aligned} \quad (2)$$

We call this the *monotonicity in minimum reference view distance*. This is also reasonable, since it is observed empirically that when both reference views are encoded at the same quality (using the same quantization parameter (QP)), the worst synthesized view distortion tends to take place at the middle view [3].

Third, we assume that the rate of distortion increases with respect to minimum reference view distance is no smaller if the current distortion is higher. Mathematically, we write:

$$\frac{\partial d_u(v^l, v^r)}{\partial v^r} = \phi(d_u(v^l, v^r)), \text{ if } |v^r - u| \leq |v^l - u| \quad (3)$$

where $\phi(\cdot)$ is a monotonically non-decreasing function. Similar assumption applies when virtual view u is closer to the left reference view. We call this the *monotonicity in reference view slope*. One example of $d_u(v^l, v^r)$ that follows this property is a linear function, in which case $\phi(y) = c$ for a constant c . Another example of $d_u(v^l, v^r)$ is an exponential function, in which case $\phi(y) = c * y$.

Using the chain rule, one can see that this assumption implies *convexity* of distortion $d_u(v^l, v^r)$ in v^r :

$$\frac{\partial^2 d_u(v^l, v^r)}{\partial v^{r2}} = \frac{\partial \phi(d_u(v^l, v^r))}{\partial d_u(v^l, v^r)} \frac{\partial d_u(v^l, v^r)}{\partial v^r} \geq 0.$$

The first term is non-negative since $\phi(y)$ is a monotonically non-decreasing function in y . The second term is also non-negative since $d_u(v^l, v^r)$ is a monotonically non-decreasing function in reference view v^r by the first assumption. Hence the second derivative is non-negative, and $d_u(v^l, v^r)$ is convex in reference view v^r . An example of distortion function $d_u(\cdot)$ satisfying the above properties is shown in Fig. 1. The assumption of convexity in distortion function is common in classical rate-distortion (RD) analysis.

For a virtual view u that itself is a camera-captured view, it can also be synthesized by a pair of left and right reference views v^l and v^r where $v^l < u < v^r$. The distortion $d_u(v^l, v^r)$ follows the same properties as discussed above. Alternatively, it can be perfectly constructed with the camera-captured view u with zero distortion. Based on the above discussion, for any virtual view u , let \mathcal{V}_u denote its selected reference view set, and the corresponding distortion is:

$$D_u(\mathcal{V}_u) = \begin{cases} d_u(v^l, v^r), & \text{if } \mathcal{V}_u = \{v^l, v^r\}, \\ 0, & \text{if } u \text{ is an camera view and } \mathcal{V}_u = \{u\}. \end{cases} \quad (4)$$

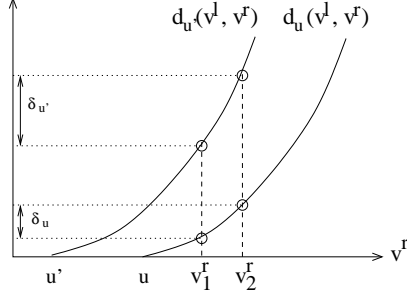


Fig. 1. Example of synthesized view distortion as function of right reference view v^r for two virtual views u and u' .

3.3. Subscription Fee Sharing

We assume that the server charges subscription fee A for streaming one video view (texture and depth). Assuming N users are close to each other in network distance, users who request the same camera view can request only one copy from the server, and share the view and the subscription fee A with each other. Let $n_v = \sum_u q_u I[v \in \mathcal{V}_u]$ be the fraction of users utilizing view v as reference, where $I[\cdot]$ is the indicator function. Thus, each user requesting view v can access view v with the subscription payment $s(v) = A/(Nn_v)$.

Based on the above discussion, given reference view selection \mathcal{V}_u , a user of virtual view u has the overall cost $c_u(\mathcal{V}_u)$ that includes two terms: synthesized view distortion and the subscription payment

$$c_u(\mathcal{V}_u) = D_u(\mathcal{V}_u) + \sum_{v \in \mathcal{V}_u} s(v). \quad (5)$$

4. AGGREGATE PERFORMANCE OPTIMIZATION

We now derive an algorithm to find a stable NE solution for users' reference view selections $\{\mathcal{V}_u^*\}$. It means that user of any virtual view u cannot deviate from the NE solution \mathcal{V}_u^* and further reduce its own cost, given that users at all other virtual views follow the NE solution $\{\mathcal{V}_u^*\}$. Mathematically, we have $c_u(\mathcal{V}_u^*) \leq c_u(\mathcal{V}_u)$ for any u given that all other virtual views follow $\{\mathcal{V}_u^*\}$. Thus, no one has incentive to unilaterally change his reference view selection. We first describe a condition for reference view selection in an NE solution. We then propose an efficient algorithm to find an NE solution $\{\mathcal{V}_u^*\}$.

4.1. Condition for an Equilibrium Solution

Given the properties of the synthesized view distortion described in Section 3.2, we state formally the following important lemma conditioning on the selection of right reference views in the equilibrium solution.

Lemma 1 In the equilibrium solution, suppose user of virtual view u chooses left and right reference views v^l and v_1^r , where $|v^l - u| > |v_1^r - u|$. Then, user of view $u' < u$ with left reference view w^l , where $w^l \leq v^l$, cannot choose right reference view v_2^r over view v_1^r , where $v_2^r \in (v_1^r, \min\{2u' - w^l, 2u - v^l\}]$.

Fig. 1 illustrates an example. A similar lemma can be written for conditioning of the selection of left reference views in the equilibrium solution as well. We now prove the above lemma as follows.

Proof of Lemma 1 We prove by contradiction. Suppose that in the equilibrium solution, a user of virtual view u' , $u' < u$, with left reference view w^l , where $w^l \leq v^l$, selects right reference view $v_2^r \in (v_1^r, \min\{2u' - w^l, 2u - v^l\}]$ over v_1^r . We consider two cases: i) $w^l = v^l$ and ii) $w^l < v^l$.

Consider first the case where $w^l = v^l$. The largest value v_2^r can take on is $2u' - v^l$. Hence

$$\begin{aligned} v_2^r &\leq 2u' - v^l \\ v_2^r - u' &\leq u' - v^l \end{aligned}$$

That means virtual view u' is always closer to right reference view v_2^r than left reference view v^l . Given $u' < u$, virtual view u is also closer to right reference view v_2^r than left reference view v^l . Further, by monotonicity in minimum reference view distance, $u' < u$ means:

$$d_{u'}(v^l, v^r) \geq d_u(v^l, v^r) \quad \forall v^r \in [v_1^r, v_2^r].$$

Consequently, by monotonicity in reference view slope, we have

$$\frac{\partial d_{u'}(v^l, v^r)}{\partial v^r} \geq \frac{\partial d_u(v^l, v^r)}{\partial v^r} \quad v^r \leq v_1^r \leq v_2^r.$$

Given $d_{u'}$ is no smaller than d_u at right reference view v_1^r and increases from v_1^r to v_2^r at a rate no smaller than d_u , we can conclude that:

$$\begin{aligned} d_{u'}(v^l, v_2^r) - d_{u'}(v^l, v_1^r) &\geq d_u(v^l, v_2^r) - d_u(v^l, v_1^r) \\ &> s(v_1^r) - s(v_2^r). \end{aligned}$$

The last inequality stems from the fact that user of virtual view u selects right reference view v_1^r over v_2^r , meaning that the drop in distortion is larger than any potential increase in subscription fee. That means user of virtual view u' can achieve a lower cost by choosing virtual view v_1^r over v_2^r . A contradiction.

We next consider the case where $w^l < v^l$. Because $w^l < v^l < u$, user of virtual view u can select w^l as left reference view for view synthesis. From the assumption of monotonicity in reference view distance, we have

$$d_u(w^l, v^r) \geq d_u(v^l, v^r) \quad \forall v^r \in [v_1^r, v_2^r].$$

Again, by monotonicity in reference view slope, we have

$$\frac{\partial d_u(w^l, v)}{\partial v} \geq \frac{\partial d_u(v^l, v)}{\partial v} \quad v^r \leq v_1^r \leq v_2^r.$$

Given the above two observations, we can conclude that

$$\begin{aligned} d_u(w^l, v_2^r) - d_u(w^l, v_1^r) &\geq d_u(v^l, v_2^r) - d_u(v^l, v_1^r) \\ &> s(v_1^r) - s(v_2^r) \end{aligned}$$

Following similar steps as in the first case, it can be shown that

$$\begin{aligned} d_{u'}(w^l, v_2^r) - d_{u'}(w^l, v_1^r) &\geq d_u(w^l, v_2^r) - d_u(w^l, v_1^r) \\ &> s(v_1^r) - s(v_2^r). \end{aligned} \quad (6)$$

This also contradicts the assumption that user of virtual view u selects v_2^r over v_1^r . Since both cases are shown to be contradictions, the lemma is proven. \square

Lemma 1 shows that when a user of virtual view u selects left and right reference views v^l and v^r , for users in any virtual view u' , $u' < u$, with left reference view w^l , $w^l \leq v^l$, they will not select right reference view w^r in the range $\delta = (v^r, \min\{2u' - w^l, 2u - v^l\}]$. Lemma 1 can help reduce the search space when we seek for the equilibrium solution.

Consider first an exhaustive search algorithm, where for each virtual view u' , it tries all possible left and right reference views w^l and w^r and chooses the optimal pair to minimize cost for view u' . If we apply lemma 1 for each virtual view u' and left reference w^l , we

first need to find out the excluded range δ and then try the remaining right reference views. In practice, determining δ can itself be computation-expensive. When $|\delta|$ is small, the saving in a reduced search space is outweighed by the computation of δ . Thus lemma 1 should only be selectively applied to speed up a search algorithm.

It is clear that when $(2u' - w^l)$ or $(2u - v^l)$ is small, $|\delta|$ will also be small. Further, u' being close to V means there are not many candidate right reference views in the first place. Thus, we set two necessary conditions before applying lemma 1 to a search algorithm: i) $2u' - w^l \geq \tau_1$, and ii) $u' \leq \tau_2$, where τ_1 and τ_2 are pre-determined parameters. We detail our search algorithm next.

4.2. Efficient Algorithm for Nash Equilibrium Solution

Using lemma 1, we describe an algorithm that finds an NE solution.

Algorithm 1 Nash Equilibrium Solution Search

```

1: Identify range  $[u^l, u^r]$  that contains all peers.
2: Initialize  $\mathcal{V}_u = \{[u^l], [u^r]\}, \forall u$ .
3: repeat
4:   for each virtual view  $u'$  with viewers do
5:     for each left reference  $w^l \in [[u^l], [u']]$  do
6:        $\delta = \emptyset$ .
7:       if  $(2u' - w^l) \geq \tau_1$  and  $u' \leq \tau_2$  then
8:         for each virtual view  $u > u'$  do
9:            $\delta = \delta \cup (v^r, \min\{2u' - w^l, 2u - v^l\})$ .
10:        end for
11:       end if
12:       for each  $w^r \in [[u^r], [u']] \setminus \delta$  do
13:         if  $c_{u'}(\{w^l, w^r\})$  is the smallest cost so far then
14:           Update  $\mathcal{V}_{u'} = \{w^l, w^r\}$ .
15:         end if
16:       end for
17:     end for
18:   end for
19: until  $\{\mathcal{V}_u\}$  is stable.

```

We first initialize a tight virtual range $[u^l, u^r]$ that contains all peers. The tightest reference views that sandwich this range are $[u^l]$ and $[u^r]$, which we use to initialize \mathcal{V}_u 's.

Then, given solution $\{\mathcal{V}_u\}$ in the last iteration, for each virtual view u' with viewers, we search its optimal reference view selection $\mathcal{V}_{u'}$ assuming that users of other views follow $\{\mathcal{V}_u\}$. Specifically, for each possible left reference w^l , we search its optimal right reference w^r that gives the lowest cost $c_{u'}(\{w^l, w^r\})$. The search range for w^r can be decreased by δ , if the two conditions $(2u' - w^l) \geq \tau_1$ and $u' \leq \tau_2$ are satisfied and lemma 1 is applied. The algorithm is repeated until the solution $\{\mathcal{V}_u\}$ is stable.

5. EXPERIMENTATION

For our experiments, we used the same synthesized view distortion function in [5]:

$$d_u(v^l, v^r) = \gamma e^{\alpha(v^r - v^l)} (e^{\beta * \min(u - v^l, v^r - u)} - 1) \quad (7)$$

which meets all the properties described in Section 3. We also used the same parameters: $\gamma = 0.06$, $\alpha = \beta = 0.2$. For available views for free viewpoint navigation, we assumed 21 captured views and 221 virtual views. We assumed each user's view distribution follows a uniform distribution, and he selects a particular view with probability $1/221$. We tested our system under different network size N and subscription fee A . We ran each simulation for 200 times and will show the average results in the following discussions.

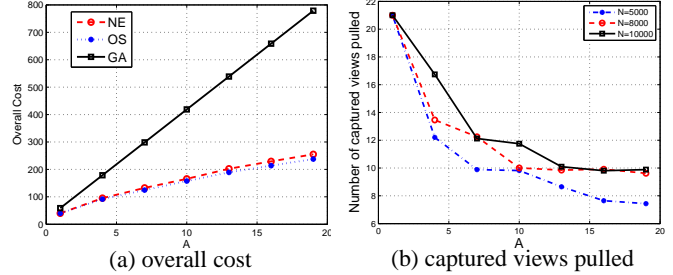


Fig. 2. Overall cost and number of captured views pulled versus subscription fee. In (a), NE, OS and GA are compared for fixed network size $N = 5000$. In (b), network size was varied $N = 5000, 8000$ and 10000 .

Fig. 2(a) shows the performance of the NE solution (NE) with the optimal reference view selection (OS) and the grouping algorithm (GA) studied in [5]. OS assumes that all users cooperatively purchase a subset of reference views $\mathcal{V}' \subseteq \mathcal{V}$, and each peer will select the tightest left/right references from \mathcal{V}' to minimize its distortion. OS aims to find the optimal \mathcal{V}' to minimize the overall cost (i.e., the total distortion and the total subscription payment of all users). Although OS gives the global optimal reference selection, it is not stable; it does not take into account users' selfish tendency to seek to reduce their own cost instead of the overall cost. In GA, all users between two neighboring camera-captured views form a group, and each group independently requests the tightest left and right reference views. The users in the same group share the subscription fee together. Fig. 2(a) shows the overall cost of the three algorithms. We first observe that GA gives the highest overall cost, since different groups do not share reference views. We also observe that NE results in a slightly higher overall cost than OS. Hence NE provides incentive for users to form stable reference selection with only a small loss of the overall performance.

Fig. 2(b) shows the total number of requested reference views selected by users via NE for different network sizes. We first observe that the number of requested reference views decreases when subscription fee A increases. This is because when A is larger, users request smaller number of reference views with more users for each reference to share the fee. We observe also that the number of requested reference views increases with increased network sizes. This is because a larger network has more users to share the cost. Since more reference views can effectively reduce users' distortions, each user can have a lower cost on average in a larger network.

Using 81 captured views in the system and setting $\tau_1 = 65$ and $\tau_2 = 70$, algorithm using lemma 1 can save more than 10% in execution time comparing to exhaustive search. Experiment results also show that algorithm with lemma 1 can reap more computation savings for large number of captured views.

6. CONCLUSION

We study the optimal reference view selection problem in a cooperative free viewpoint streaming system, where a user can simultaneously belong to two groups and share the cost of a single reference with each group separately. To study a stable NE reference selection, we propose an efficient search algorithm, leveraging on the properties of the synthesized video distortion. Simulation results show that a stable NE solution only slightly increases the overall cost, when compared to the unstable optimal reference selection that gives the lowest overall cost. Furthermore, a larger network has the ability to request more reference views to reduce users' distortion without much increase of the subscription fees shared by each user.

7. REFERENCES

- [1] M. Tanimoto, M. P. Tehrani, T. Fujii, and T. Yendo, "Free-viewpoint TV," in *IEEE Signal Processing Magazine*, January 2011, vol. 28, no.1.
- [2] W. Mark, L. McMillan, and G. Bishop, "Post-rendering 3D warping," in *Symposium on Interactive 3D Graphics*, New York, NY, April 1997.
- [3] P. Merkle, Y. Morvan, A. Smolic, D. Farin, K. Muller, P. de With, and T. Wiegand, "The effects of multiview depth video compression on multiview rendering," in *Signal Processing: Image Communication*, 2009, vol. 24, pp. 73–88.
- [4] G. Cheung, V. Velisavljevic, and A. Ortega, "On dependent bit allocation for multiview image coding with depth-image-based rendering," in *IEEE Transactions on Image Processing*, March 2011, vol. 20, no.11, pp. 3179–3194.
- [5] D. Ren, S.-H. G. Chan, G. Cheung, V. Zhao, and P. Frossard, "Collaborative P2P streaming of interactive live free viewpoint video," in *arXiv.org*, November 2012, <http://arxiv.org/abs/1211.4767>.
- [6] G. Owen, *Game Theory*, Academic Press, 3rd edition, 1995.
- [7] P. Merkle, C. Bartnik, K. Muller, D. Marpe, and T. Weigand, "3D video: Depth coding based on inter-component prediction of block partitions," in *2010 Picture Coding Symposium*, Krakow, Poland, May 2012.
- [8] I. Daribo, D. Florencio, and G. Cheung, "Arbitrarily shaped sub-block motion prediction in texture map compression using depth information," in *2010 Picture Coding Symposium*, Krakow, Poland, May 2012.
- [9] E. Kurutepe, M. R. Civanlar, and A. M. Tekalp, "Client-driven selective streaming of multiview video for interactive 3DTV," in *IEEE Transactions on Circuits and Systems for Video Technology*, November 2007, vol. 17, no.11, pp. 1558–1565.
- [10] H. Huang, B. Zhang, G. Chan, G. Cheung, and P. Frossard, "Coding and replication co-design for interactive multiview video streaming," in *mini-conference in IEEE INFOCOM*, Orlando, FL, March 2012.
- [11] H. Huang, B. Zhang, G. Chan, G. Cheung, and P. Frossard, "Distributed content replication for multiple movies in interactive multiview video streaming," in *19th International Packet Video Workshop*, Munich, Germany, May 2012.
- [12] B. Macchiavello, C. Dorea, M. Hung, G. Cheung, and W. t. Tan, "Reference frame selection for loss-resilient depth map coding in multiview video conferencing," in *IS&T/SPIE Visual Information Processing and Communication Conference*, Burlingame, CA, January 2012.
- [13] B. Macchiavello, C. Dorea, M. Hung, G. Cheung, and W. t. Tan, "Reference frame selection for loss-resilient texture & depth map coding in multiview video conferencing," in *IEEE International Conference on Image Processing*, Orlando, FL, September 2012.
- [14] R. Kumar, Y. Liu, and K. Ross, "Stochastic fluid theory for P2P streaming systems," in *IEEE INFOCOM*, Anchorage, Alaska, May 2007.
- [15] X. Hei, C. Liang, J. Liang, Y. Liu, and K. W. Ross, "A measurement study of a large-scale P2P iptv system," *IEEE Transactions on Multimedia*, vol. 9, no. 8, pp. 1672–1687, Dec. 2007.
- [16] Y. Liu, Y. Guo, and Chao Liang, "A survey on Peer-to-Peer video streaming systems," *Journal of Peer-to-Peer Networking and Applications*, by Springer New York, Feb. 2008.
- [17] K. Oh, S. Yea, and Y.-S. Ho, "Hole-filling method using depth based inpainting for view synthesis in free viewpoint television and 3D video," in *27th Picture Coding Symposium*, Chicago, IL, May 2009.
- [18] I. Daribo and B. Pesquet-Popescu, "Depth-aided image inpainting for novel view synthesis," in *IEEE International Workshop on Multimedia and Signal Processing*, Saint-Malo, France, October 2010.