

# Inexpensive High Dynamic Range Video for Large Scale Security and Surveillance

Stephen Mangiat and Jerry Gibson  
Electrical and Computer Engineering  
University of California, Santa Barbara, CA 93106  
Email: {smangiat, gibson}@ece.ucsb.edu

**Abstract**—We describe a new method for High Dynamic Range (HDR) Video using alternating exposures that adds no additional cost or bandwidth requirements to individual IP cameras, making it suitable for large scale security and surveillance systems. Sufficient dynamic range is crucial to the efficacy of a surveillance system, as saturated pixels mean a camera can no longer “see” its surrounding environment. High costs associated with hardware for improved dynamic range make them unsuitable for very large networks with hundreds or even thousands of cameras. We outline a scalable software method that uses post-processing to combine the information in adjacent frames of a video sequence captured with alternating short and long exposures. In particular, we introduce a novel bi-directional motion estimation module that utilizes block-based motion vectors to register frames with large differences in global brightness and fast local motion within saturated regions. An HDR post-processing solution can be deployed at a central location to process individual camera streams on an “as needed” basis, removing additional costs at the device-end. Furthermore, cameras continue to transmit low dynamic range frames, so there is no additional bandwidth requirement. Results show significant gains in video quality for inexpensive cameras when exposed to brightness variations common in security and surveillance.

## I. INTRODUCTION

Surveillance and security cameras are vital for the protection of borders, military bases, security checkpoints, and airports, as well as countless businesses and homes. The usefulness of a camera for these tasks is strictly determined by video quality, which is ultimately balanced against costs. In a surveillance or security application, the pixel resolution, field of view, and dynamic range are paramount. Frame rate and temporal fidelity are useful for automated activity detection and object tracking, yet they are secondary to the camera’s main objective to “see” its surrounding environment.

A common surveillance task is the identification of people of interest. Here, a camera needs an unobstructed view of the person’s face or other distinguishing features with sufficient clarity. Advances in high definition sensor technology have significantly improved this clarity. Still, obtaining adequate views of a scene, or “covering all angles,” necessitates the deployment of many cameras, often on the order of hundreds or thousands in a single network. Such a large deployment requires significant financial investment, and thus cost-effective IP cameras are needed [11].

This work has been supported by Huawei Technologies, Co. Ltd., Santa Clara, CA

Camera dynamic range is equally important in this scenario. Security and surveillance cameras are often placed outdoors or near entrances to buildings, exposing them to extreme variations in brightness. Most cameras capture 8 bits per color channel (256 levels), whereas an outdoor sunlit scene might require more than 10,000 levels. Auto-exposure algorithms attempt to minimize the resultant pixel saturation, yet they fail to correctly expose the entire frame. High dynamic range (HDR) video aims to accurately record scenes with brightness variations beyond the capabilities of a typical camera sensor.



Fig. 1. (a) Low dynamic range frame captured by traditional auto-exposure (b) High dynamic range (HDR) frame created using the alternating exposure technique — note improved clarity and details in the foreground.

Limited dynamic range means that inexpensive cameras cannot “see” everything within their field of view at the same time. This hinders the identification of people or objects of interest, as well as the general understanding of a scene. Such a scenario is illustrated in Fig. 1 (a), which shows a building entrance captured by a single, low dynamic range exposure. Here, the difference in brightness between the outside and inside of the building is so large that it is impossible for the camera to adequately expose both regions simultaneously. There is significant pixel saturation not only outside, but also under and around the chairs in the foreground. Any bag or object placed under one of these chairs would not be captured by the camera, despite being well within the camera’s field of view. It is also important to note that the image shown here is not a raw image, and it has been processed to enhance local contrast.

Most HDR video methods include a way to obtain multiple exposures of a scene, using specialized hardware or software [10]. Hardware modifications such as beam splitters, multiple sensors, or spatially varying optical filters drive up costs

considerably. The desire to keep the cost of each camera low thus motivates a software-based post-processing solution, as described in Sec. II. Transmission of an alternating exposure video sequence is discussed in Sec. III, followed by an overview of HDR post-processing in Sec. IV. In Sec. V, we introduce a novel bi-directional motion estimation method that is crucial for the elimination of ghosting and the creation of multiple exposures at each time instant. The results of two sample frames are discussed in Sec. VI. Finally, Sec. VII outlines some conclusions and future work.

## II. HDR VIDEO CAPTURE

The benefits of increased dynamic range are shown in Fig. 1 (b). Here, the pixel saturation surrounding the chairs in the foreground is eliminated and new details are revealed. This image was created by alternating the camera’s exposure between a short and long exposure, and combining the information in adjacent frames. As opposed to still images, video poses significant challenges due to motion, which will cause ghosting in the HDR output if it is not compensated. Furthermore, occlusions and other limits of frame registration ultimately mean that there is a tradeoff between temporal fidelity and dynamic range, though filtering can reduce the effect of artifacts [8].

Despite some loss in temporal fidelity, which is less crucial for security and surveillance, there are a number of advantages to using an alternating exposure approach. First, it requires no hardware modifications and can be performed on video captured from very inexpensive cameras. IP cameras (cameras that transmit their data through a network connection to a centralized server) can be easily programmed to capture scenes with alternating exposures when needed. This data can also be processed on an “as needed” basis at a central location with access to much larger computing resources and power. In this way, there is no additional cost at the device-end for HDR capability.

Typical video cameras use a single exposure setting that is adapted according to the statistics of each frame. Since the dynamic range of the scene is usually much larger than that of the camera, this auto-gain control algorithm attempts to minimize the number of saturated pixels. In order to extend dynamic range, we adapt two exposures (short and long) in real-time, as in [6]. The camera cycles between these two shutter speeds in alternating frames. Our goal here is to maximize the long exposure and minimize the short exposure, thus maximizing the dynamic range expansion, while maintaining enough non-saturated pixels to adequately register adjacent frames. A more detailed description of our “dual-exposure” algorithm can be found in [8]. Instead of minimizing the number of saturated pixels in each frame, the number of saturated pixels is kept at a small percentage (typically between 20-30%). Figure 2 shows a sample long and short exposure, as well as corresponding single exposure and HDR frames.

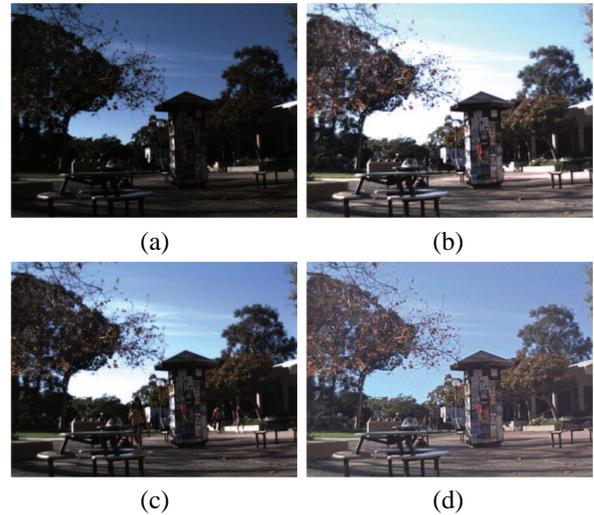


Fig. 2. Dual-Exposure Control: (a) Short Exposure (b) Long Exposure (c) Standard Auto-Exposure: Saturation causes a white sky and shadows obscure details (d) HDR Output: Enhanced colors and local contrast, without saturation (Images best viewed in color)

## III. TRANSMISSION

Another advantage of an alternating-exposure technique concerns a main technological hurdle for many large-scale surveillance camera networks: bandwidth. Cameras generate an enormous amount of data, and high dynamic range increases the number of bits per pixel. In order to view an HDR video, this high bit-depth information must be mapped back into displayable range (tone-mapped) for low dynamic range displays [9]. However, placing this processing at the camera itself will increase the costs and complexity at the device-end, limiting scalability. As such, the camera must *transmit HDR information* to a central server or “cloud.”

A state-of-the-art video camera with HDR sensors may generate up to 42 GB per minute of data, without compression [2]. For comparison, a high-definition 1080p camera with standard dynamic range has a bandwidth of less than 11 GB per minute without compression. Advanced compression techniques will reduce these numbers greatly, yet the HDR video will still represent a significant increase in bandwidth compared to low dynamic range.

However, data generated from a low dynamic range camera capturing alternating short and long exposures is still low dynamic range before it is processed. Due to the temporal subsampling of the dynamic range, the bits per pixel is not increased. Since there are large global brightness variations between adjacent frames, a compression scheme must encode the even and odd frames separately, which may decrease compression efficiency. Yet, an increased number of saturated pixels within each stream means there is less high frequency information to encode. As such, an alternating exposure HDR method represents a negligible change to the required bandwidth for transmission.

#### IV. HDR POST-PROCESSING

Given a sequence of alternating exposures provided by a dual-exposure algorithm, the task is to utilize neighboring frames to predict a second exposure for each time instant. Ideally, this prediction should represent *exactly* the same scene as the current frame, though this is hindered due to occlusions and non-overlapping regions. Still, HDR post-processing provides very useful results. An overview of our processing pipeline is found in Fig. 3.

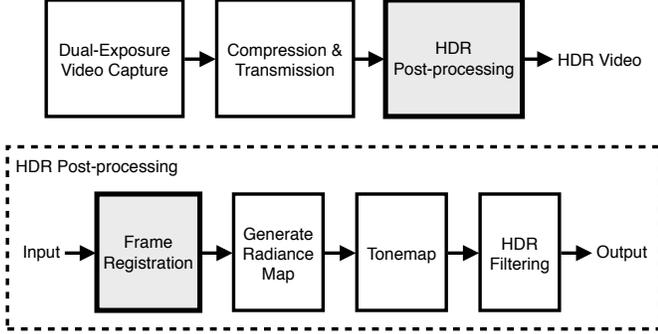


Fig. 3. High Dynamic Range Video Post-processing Overview: A novel frame registration technique for alternating exposures is outlined in Sec. V.

The first step is adjacent frame registration, and details of our approach are found in Sec. V. Following registration, the current frame and prediction are combined to form a high dynamic range radiance map using the camera response function [5], which can be estimated during the manufacturing stage. Given pixel values  $Z_{ij}$  and shutter times  $\Delta t_j$ , one can recover a high dynamic range *radiance map* using

$$\ln E_i = \frac{\sum_{j=1}^P w(Z_{ij})(g(Z_{ij}) - \ln \Delta t_j)}{\sum_{j=1}^P w(Z_{ij})}, \quad (1)$$

where  $w(Z_{ij})$  is a weighting function,  $i$  is the spatial index,  $j$  is the frame index, and  $g$  represents the log of the camera response curve [5]. Once the HDR radiance map is calculated, it must be “tone mapped” back into displayable range. We use the method described in [9], which has global and local normalization and uses a dodging and burning technique to minimize halo effects.

The result is an HDR version of the current frame that may be vulnerable to blocking and other artifacts due to the limitations of registration. Artifacts are addressed within saturated regions using a pixel-wise refinement step (see Sec. V-A), and remaining artifacts may be filtered prior to output [7]. In [8], we describe a “High Dynamic Range (HDR) Filter” that can mitigate these artifacts for perceptually pleasing HDR video without exact registration. This filter builds upon the bilateral filter to smooth frames while maintaining important edges. Additionally, the filter strength locally adapts to corresponding motion vectors. Since regions with poor registration generally correspond to faster motion, smoothing here can eliminate noticeable artifacts without degradation in perceptual quality.

#### V. FRAME REGISTRATION

In order to generate an HDR output with the same framerate as an input sequence of alternating exposures, two exposures must be available at every time instant. This requires accurate motion estimation (ME) to determine pixel correspondences between adjacent frames. In addition, this process has unique challenges due to the severe illumination change between frames and the resultant saturated pixels. The “HDR stitching” method in [6] used gradient-based optical flow ME, while the method presented here uses a block-based approach, extending the work of [7].

The first step in our frame registration approach is to calculate the forward and backward motion vectors for the current frame with respect to the previous and next frames. Since the brightness constancy assumption is violated between these frames, we must boost the short exposure,  $Z_s$ , to match the long exposure,  $Z_l$ , using

$$\hat{Z}_l = g^{-1}(g(Z_s) - \ln \Delta t_s + \ln \Delta t_l), \quad (2)$$

where  $\Delta t_s$  and  $\Delta t_l$  are the short and long exposure times, and  $g^{-1}$  is the inverse camera response function modeled by an exponential curve. We then use the H.264 JM Reference software with Enhanced Predictive Zonal Search (EPZS), a  $16 \times 16$  block size, and Sum of Absolute Differences (SAD) in both luma and chroma components to estimate the forward and backward motion vector fields [1]. The two motion fields are combined by selecting the motion vector with minimum SAD, and labels are stored to reference either the previous or next frame for each block.

##### A. Determining Poorly Registered Pixels

Due to pixel saturation, some information needed for forward/backward motion estimation is lost, producing artifacts. Therefore following forward/backward ME, we next identify poorly registered regions that must be improved using *bi-directional* motion estimation. We can locate registration errors on an RGB pixel-wise basis, and use these to assess registration quality for each block.

First, a pixel is designated as being “flipped” if it disobeys the brightness monotonicity assumption, i.e. it is brighter in the shorter exposure than it is in the longer exposure. Secondly, we identify pixels where the predicted *radiance* is poor. The absolute difference between the radiances given by the predicted pixel and the pixel in the current frame is compared to a threshold (only for pixels that are non-saturated in both frames). Finally, using the camera response curve it is possible to determine the minimum brightness in a short exposure that will over-saturate in the long exposure ( $Z_{s*} = g^{-1}(g(Z_{max}) - \ln \Delta t_l + \ln \Delta t_s)$ ), as well as the maximum brightness in the long exposure that will under-saturate in the short exposure ( $Z_{l*} = g^{-1}(g(Z_{min}) - \ln \Delta t_s + \ln \Delta t_l)$ ). For instance, if the current frame is a long exposure, we can locate pixels that are saturated in the current frame whose predicted values are less than the threshold  $Z_{s*}$ . Pixels are labeled as “bad” if any of these criteria are met in at least one color channel.

## B. Determining Poorly Registered Blocks

In [7], blocks were labeled as “saturated” if the number of saturated pixels within the block was greater than 50% of the entire block. However, this method only identifies a subset of potentially mis-registered blocks, as blocks with little texture may be assigned incorrect motion vectors despite having good matches. As such, we expand this “saturated” classification to include blocks in the current frame with standard deviation less than a threshold, and blocks in the prediction frame with standard deviation less than a threshold. The saturated pixel threshold is also adjusted to 60% of the entire block.

All blocks labeled as “saturated” will be addressed using bi-directional motion estimation. Additionally, we identify a subset of blocks that may contain sufficient texture for matching, yet we cannot trust that their motion vectors represent the true motion. These blocks, labeled as “unreliable”, typically appear in regions where objects are partially occluded. Their corresponding SAD cost may be quite low, so these blocks are not replaced if they are not classified as “saturated.” However, it is necessary to mark them as “unreliable” since the motion vectors of reliable blocks will be utilized during bi-directional prediction. A block is labeled “unreliable” if the number of bad pixels within the block is greater than a threshold (See Sec. V-A), or the length of its motion vector is greater than a chosen threshold. Very large MVs (greater than 60 pixels at 30 fps) are most likely remnants of inaccurate motion estimation.

## C. Bi-directional Motion Estimation

Once blocks are labeled as “saturated” or “unreliable”, the previous and next frames are prepared for block-based bi-directional motion estimation. Since this involves calculating the SAD and mean absolute difference (MAD) between blocks in the previous and next frames, it is again imperative that these frames have the same global brightness. Despite having the same classification as either short or long exposure frames, they might have slightly different exposure times due to the dynamic adjustment of shutter speed. Consequently, the frame with the shorter exposure time is boosted to match the brightness of the longer exposure time, as in forward/backward motion estimation.

1) *Zero Motion*: In security and surveillance applications, the camera is often stationary or panning. A global panning motion is captured well by a 2D homography, which may be estimated using block motion vectors. For stationary cameras, it is possible that much of the frame will have no motion. As such, the first step of bi-directional ME is to check every block in the frame for zero motion. We check every block, instead of only blocks labeled as “saturated,” since whether or not a block has zero motion is an important distinction for HDR filtering [8].

For a given block, we first calculate the SAD between co-located blocks in the previous and next frames and compare it to a threshold. However, it is important to check whether the co-located block has sufficient texture for matching in the prediction frame (either the previous or next frame, depending on which provides the greatest dynamic range expansion).

If it is saturated or too smooth, we cannot trust that a zero motion vector is accurate. Still, a *radiance based* background subtraction model can help reduce ambiguity. This model can be calculated periodically or adaptively using one of several methods, such as the median of previous frames [4]. Radiances predicted by non-saturated pixels in the current frame and adjacent frames are compared to the model, and pixels with a radiance difference below a threshold are labeled as “background.”

At this stage, we also note blocks that have low zero motion SADs, but their co-located blocks do not have sufficient texture, and neither the blocks in the current frame or co-located blocks in the previous/next frames are labeled as “background.” These blocks are likely saturated in both the long and short exposures. Since there is little information for motion estimation here, they are saved for last during later stages so more neighboring information is available. Yet if a block does meet all criteria, we treat its corresponding RGB block as a candidate, and check for “bad pixels” with respect to the RGB block in the current frame, as described in Sec. V-A.

The total number of bad pixels in the candidate RGB block,  $n_{\text{bad}}$ , is used as an additional cost factor for the candidate zero motion vector block with

$$\text{cost}_{\text{total}} = \text{SAD} + \lambda n_{\text{bad}}, \quad (3)$$

where  $\lambda$  is an empirically chosen constant. Finally, if  $\text{cost}_{\text{total}}$  is below a chosen threshold, the current block is assigned the zero motion vector, the co-located block in the prediction frame is used as the final prediction, and the block is removed from the “saturated” or “unreliable” lists (if necessary). Furthermore, this block is assigned a new reference label to signify that both the previous and next frames are locally valid for prediction.

2) *Non-zero Motion*: Following zero motion prediction, a novel bi-directional motion search is used to improve the predictions of any remaining “saturated” blocks. This process is completed over multiple passes, and utilizes the medoids of neighboring reliable motion vectors to initialize each search. For a given “saturated” block, the first step is to count the number of reliable neighbors (maximum of eight possible neighbors). If the number of good neighbors is greater than or equal to three and at least one of them shares a border with the current block, then the current block will be processed. If there are not yet enough valid neighbors, then it is saved for a later pass. The MVs of “unreliable” blocks (see Sec. V-B) are only used when there are not enough reliable neighbors to process the entire frame.

When a block has a sufficient number of reliable neighbors, we check the labels of these neighbors and count the number of blocks from the previous frame,  $n_p$ , and the number of blocks from the next frame,  $n_n$ . Accordingly, we define the number of neighbors that share a border with the current block as  $n_{b_p}$  and  $n_{b_n}$ . These counts are used to determine the local prediction frame for the current block, by choosing the maximum of  $n_p$  and  $n_n$ , or  $n_{b_p}$  and  $n_{b_n}$  (when  $n_p = n_n$ ). If  $n_{b_p}$  and

$n_{b_n}$  are also equal, then the prediction frame is chosen by the label assigned during forward/backward ME. Furthermore, the predicted motion vector,  $MV_{\text{pred}}$ , is the medoid of neighboring reliable blocks from this predicted frame only. However, if  $n_p$  was equal to  $n_n$ , then  $MV_{\text{pred}}$  is the medoid of all neighboring reliable blocks.

A bi-directional motion search is now centered about  $\mathbf{p} + MV_{\text{pred}}$ , where  $\mathbf{p}$  represents the indices of the current block (fast motion search algorithms may be used for reduced complexity). The cost to be minimized includes two main factors: the mean absolute difference between blocks in the previous and next frames ( $MAD_{\text{bidirect}}$ ) and the *boundary match* ( $MAD_{\text{bound}}$ ) between the candidate blocks and the neighboring reliable blocks in the predicted frame. The boundary match algorithm is an important part of macro-block recovery techniques and error resiliency [3], and it works well under the assumption that video frames are smooth at block boundaries. This assumption is not always valid, so it is used here in conjunction with the bidirectional MAD. The *mean* absolute difference is used instead of SAD since the number of reliable boundaries,  $n_b$ , varies from block to block. The relative weighting between these costs is also varied, with

$$\text{cost}_{\text{MAD}} = (1 - \alpha n_b)MAD_{\text{bidirect}} + \alpha n_b MAD_{\text{bound}}. \quad (4)$$

In this way, the importance of the boundary match increases with the number of reliable boundaries. In our tests, we chose  $\alpha$  to be .15. Figure 4 illustrates the boundary matching region with a block size of  $16 \times 16$ , as well the MV predicted by the medoid of neighboring reliable MVs. For increased fidelity, a smaller  $8 \times 8$  block size may also be used here.

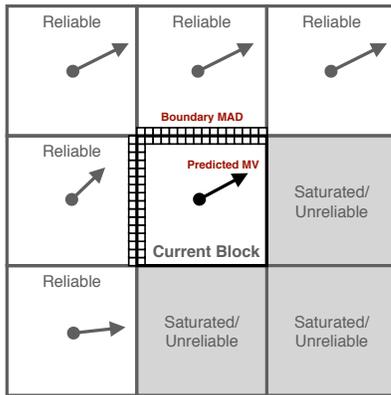


Fig. 4. Bi-directional motion search: Boundary matching and motion vector prediction (used to initialize the search) using reliable neighboring blocks.

In addition to costs from boundary matching and bidirectional MAD, we can check pixels in each RGB candidate block with the corresponding RGB block in the current frame, as described in Sec. V-A. Finally, we add one additional cost term as in [7] in order to promote motion vector field smoothness: the distance between  $MV_{\text{pred}}$  and the candidate MV. The total cost is now

$$\text{cost}_{\text{total}} = \text{cost}_{\text{MAD}} + \lambda_1 n_{\text{bad}} + \lambda_2 \|MV_{\text{pred}} - MV_{\text{cand}}\|, \quad (5)$$

where  $\lambda_1$  and  $\lambda_2$  are empirically chosen constants.

It is important to note that even when a block is labeled as “saturated,” it is not certain that the motion vector assigned by forward/backward ME is incorrect. In fact, for some saturated blocks the bi-directional prediction will be worse. For instance, if a block has a zero motion vector with respect to the previous frame and it is occluded in the next frame, the match for zero motion directly between the previous and next frames will be very poor. This means that bi-directional ME should only *attempt* to replace the forward/backward prediction. In this way, the bi-directional prediction will not be used when the associated cost is too high, or  $\text{cost}_{\text{total}} > \text{cost}_{\text{max}}$ . For most blocks, we set  $\text{cost}_{\text{max}}$  to a fairly low value, and thus require the search to find a good match. However, there are some blocks for which we raise  $\text{cost}_{\text{max}}$  to increase the likelihood that the new prediction will be used. These include blocks that have been labeled both “saturated” and “unreliable,” and blocks with a large difference between the MV assigned by forward/backward ME and the new candidate MV. If this difference is extremely large, then it is likely that the original MV is incorrect and thus appropriate to increase  $\text{cost}_{\text{max}}$ . Conversely, if this difference is very small, then it is likely that the original MV is correct and may perform better than the new bi-directional prediction.

## VI. RESULTS

To test our HDR post-processing methods, we captured sequences of alternating short and long exposures (30 fps and  $640 \times 480$  resolution) using a Point Grey Research Firefly camera<sup>1</sup>. Furthermore, we often recorded several frames with a single exposure level before the HDR mode was engaged, to allow a comparison such as that in Fig. 1.

Sample input frames from two test videos are shown in Fig. 5 (a) and (e). These frames are both short exposures, exhibiting local motion across regions with a significant number of undersaturated pixels. In fact, much of the frames are unusable for direct motion estimation. This is due to the large brightness variations found in both scenes. In the top sequence, the camera sits in the shade as cars drive by under direct sunlight. Using a low dynamic range camera, it is impossible to adequately expose the foreground and background simultaneously. Information that may prove important in a surveillance scenario, such as the car’s license plate or distinguishing features, might ultimately be lost due to pixel saturation, even if the image resolution is sufficiently high. Similarly, the bright sunlight passing through the windows in Fig. 5 (e) makes it impossible to correctly expose both indoors and outdoors, a common problem near building entrances.

The importance of the bi-directional motion estimation process (Sec. V-C) is illustrated in Fig. 5 (b)-(c) and (f)-(g). First, Fig. 5 (b) and (f) show long exposure predictions of the current frames created from adjacent frames *without* bi-directional prediction. In Fig. 5 (b), registration is very poor underneath the yellow car, a region where the input frame is

<sup>1</sup>For videos, please visit <http://vivonets.ece.ucsb.edu/HDR.html>



Fig. 5. High Dynamic Range Video Results (Images best viewed in color)

completely saturated. Similarly, the registration quality is poor throughout the foreground in Fig. 5 (f). Furthermore, there are visible artifacts across the walls in the background due to their minimal texture, which led to incorrect MVs.

The predicted long exposure frames *after* bi-directional motion estimation are shown in Fig. 5 (c) and (g). The registration quality throughout the areas saturated in the current frames is improved significantly. The zero-motion stage described in Sec. V-C1 has fixed the registration errors found across the walls in Fig. 5 (f). Furthermore, the process has performed well here despite the complex motion of the man walking behind an occluding object. The final HDR outputs are shown in Fig. 5 (d) and (g). Most of the frames are now exposed correctly, with good color information. In fact, regions that were undersaturated in the original frames now appear brighter than in the predicted long exposures.

## VII. CONCLUSIONS & FUTURE WORK

High dynamic range video will be an important component of future security and surveillance networks, even when costs must be limited at each camera. We have outlined a system that utilizes alternating exposures and post-processing to expand the dynamic range of inexpensive camera sensors, with negligible cost and bandwidth increases at the device-end. Furthermore, we have proposed a new bi-directional motion estimation algorithm that can register complex local motion found within saturated image regions.

The post-processing described here achieves important gains in dynamic range with respect to a single exposure. The trade-off is some loss of temporal fidelity, yet this is secondary for security and surveillance videos. Still, post-processing might ultimately be implemented in a scalable fashion. Significant computational complexity might only be devoted to frames

with important activity, in order to create the highest quality outputs when needed. Future work might extend the frame registration process to include moving cameras that exhibit global motion, as well as study the effects of compression on output quality. Furthermore, a number of complexity reduction techniques including parallel processing may be explored.

## REFERENCES

- [1] "H.264/AVC JM reference software," <http://iphome.hhi.de/suehring/tml/>, 2008.
- [2] A. Chalmers, "Surgeons, CCTV & TV football gain from new video technology that banishes shadows and flare," [http://www2.warwick.ac.uk/newsandevents/pressreleases/surgeons\\_cctv\\_tv/](http://www2.warwick.ac.uk/newsandevents/pressreleases/surgeons_cctv_tv/), Jan. 2011.
- [3] Y. Chen, Y. Hu, O. Au, H. Li, and C. W. Chen, "Video error concealment using spatio-temporal boundary matching and partial differential equation," *Multimedia, IEEE Transactions on*, vol. 10, no. 1, pp. 2–15, Jan. 2008.
- [4] S.-C. S. Cheung and C. Kamath, "Robust techniques for background subtraction in urban traffic video," S. Panchanathan and B. Vasudev, Eds., vol. 5308, no. 1. SPIE, 2004, pp. 881–892.
- [5] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *SIGGRAPH '97*. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1997, pp. 369–378.
- [6] S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, "High dynamic range video," in *ACM SIGGRAPH*, New York, NY, USA, 2003, pp. 319–325.
- [7] S. Mangiat and J. Gibson, "High dynamic range video with ghost removal," in *SPIE Optical Engineering & Applications*, 2010.
- [8] S. Mangiat and J. Gibson, "Spatially adaptive filtering for registration artifact removal in hdr video," in *IEEE International Conference on Image Processing (ICIP)*, sep. 2011.
- [9] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Transactions on Graphics*, vol. 21, no. 3, pp. 267–276, 2002.
- [10] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec, *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2005.
- [11] M.-J. Yang, J. Y. Tham, D. Wu, and K. H. Goh, "Cost effective ip camera for video surveillance," in *Industrial Electronics and Applications, 2009. ICIEA 2009. 4th IEEE Conference on*, May 2009, pp. 2432–2435.