E-ISSN: 2321–9637
Volume 2, Issue 1, January 2014
International Journal of Research in Advent Technology
Available Online at: http://www.ijrat.org

AUTOMATIC VIEW SYNTHESIS FROM STEREOSCOPIC 3D BY IMAGE DOMAIN WARPING

Mrs. Mayuri T. Deshmukh
Assistant Professor
S.S.B.T's C.O.E.T. Bambhori.
Email id:mayuri.deshmukh08@gmail.com

Miss.Ashwini T. Sharnagat
ME Ist Digital Electronic
S.S.B.T's C.O.E.T. Bambhori.
Email id:aashwini179@gmail.com

ABSTARCT- This paper focuses on the automatic view synthesis by image domain warping is presented that automatically synthesizes new views directly from S3D video and functions completely. Nowadays, stereoscopic 3D (S3D) cinema is already conventional and almost all new display devices for the home support S3D content. S3D giving out communications to the home is already established partly in the form of 3D Blu-ray discs, video on demand services, or television channels. The need to wear glasses is, however, often considered as an obstacle, which hinders broader acceptance of this technology in the home. Multiviewautosterescopic displays make possible a glasses free perception of S3D content for several observers simultaneously and support head motion parallax in a limited range. To verify the entire system the result is being observed on 3D television, 3D cinema's for view the video allow a glasses free perception.

Index Terms—Three dimensional TV, auto stereoscopic displays, 3D Blu-ray discs, head motion parallax

1. INTRODUCTION

STEREOSCOPIC 3D (S3D) cinema and television are in the process of changing the scenery of entertainment. Primarily responsible for the change is the fact that technologies ranging from 3D content creation, to data compression and transmission, to 3D display devices are progressively improving and adapted to enable a rich and higher quality 3D experience. However, the necessity to wear glasses is often regarded as a main obstacle of today's conventional stereoscopic 3D display systems. Multi-view auto stereoscopic displays (MAD) overcome this problem. They allow glasses free stereo viewing by emitting several images at the same time.

Stereoscopic 3D can expand users' experiences beyond traditional2D-TV broadcasting by contribution programs with depth idea of the observed scenes. IN fact, 3D has been successfully commercialized as stereo movies, such as those by IMAX, for people to watch in the cinema, using special devices. Given that the popularity of 3D programs has dramatically increased, 3D-TV has been known as a possible break through for conventional TV technologies to satisfy the coming need for watching 3D programs at home. Typical MADs which are on the market today require 8-views, 9-views or even 28-views as input. Because of the different number of input views required by different MADs, no unique display format exists for such displays**.** According to the formats involved in the distribution chain (Fig.1.), the transmission format has to enable such a decoupling. Hence, a good transmission format has to fulfil the following requirements

• An automatic conversion from production to transmission format has to be possible. For live broadcast applications also real-time conversion is required.

E-ISSN: 2321–9637
Volume 2, Issue 1, January 2014
International Journal of Research in Advent Technology
Available Online at: http://www.ijrat.org

• The transmission format has to allow an automatic and real-time conversion into any particular N-view display format.

• The transmission format has to be well compressible to save transmission band width. nowadays, professional and consumer 3D content production is dominated by S3D content, i.e. 2-view content which is watchable on stereoscopic displays. It is believed that S3D as aproduction format will dominate over years[Nikolce and Stefanoski,2013]

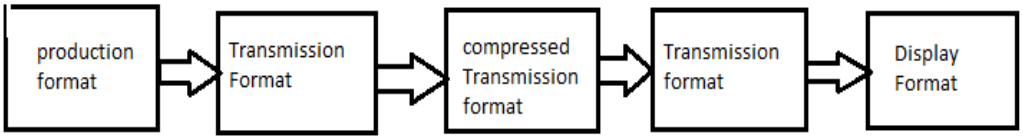

Fig.1. High level view on the usual format conversions required from content Production to display

Research communities and standardization bodies continue to investigate formats which are well compressible and enable efficient generation of novel views as required by MADs. Such formats can be divided into two classes: video only formats like S3D and multiview video in general, and depth enhanced formats like single view video plus depth and multiview video plus depth. The per-pixel depth information included in the depth enhanced formats provides information on the depth structure of the 3D scene. Such depth data can be used for view synthesis by depth-image based rendering (DIBR) [Masayuki T, Zhao Y, and C. Zhu, 2012], [Fehn C, 2004].

However, high quality view synthesis with DIBR requires high quality depth data. There exist stereo algorithms which can automatically compute depth maps from stereo images, or depth sensors which can capture depth usually of too little accuracy to allow a high quality synthesis as required e.g. in professional content productions. Thus, today highly accurate depth maps are usually generated in a semiautomatic process, where stereo algorithms or depth sensors are used to estimate or capture initial depth maps which are then improved in an interactive process. The most simple and cost efficient conversion from S3D As a production format to a transmission format consists of conducting only low level conversion steps (like colour space, bit depth, frame-rate, or image resolution conversions). Such a transmission format can be well compressed and is compatible to the existing S3D distribution and stereoscopic display infrastructure. Thus, using a transmission format without supplementary depth data prevents an increase of content production and distribution costs. The presented synthesis method can be used at the decoder side to synthesize new views directly from transmitted S3D content. for real-time view synthesis is presented and analysed.

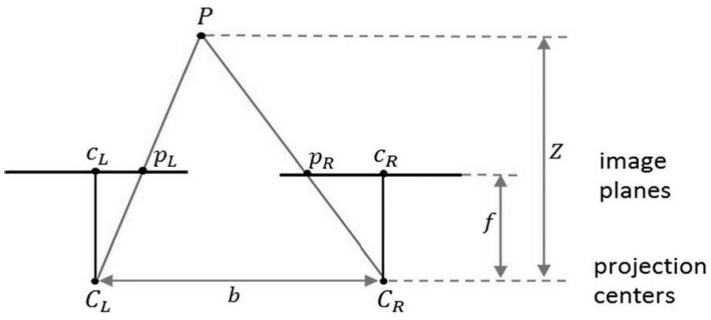

Fig.2 Stereo pinhole camera model illustrating the projection of a 3D point pinto the image planes of both cameras.

It shows a bird's-eye view of a pin-hole model, parallel stereo camera setup. Projection centers of both cameras are located at positions *CL* and *CR* at a baseline distance of *b* and both cameras have a focal length of f

E-ISSN: 2321–9637
Volume 2, Issue 1, January 2014
International Journal of Research in Advent Technology
Available Online at: http://www.ijrat.org

projecting a 3D point *P*, which is located at a distance *Z* from the projection centers, into the respective**.**projection planes gives the projected points *pL* and *pR*. Hence, the projected points have an image disparity of

$$d = (pL - cL) - (pR - cR) = \frac{F_z}{b} \quad (1)$$

Obviously, a synthesis of a new view at a position

$$C \text{ new} = (1 - \lambda)\, C_L + \lambda C_R \quad (2)$$

Corresponds to having a baseline distance

$$b_{new} = \lambda b \quad (3)$$

with respect to the left view. Hence, that would give disparities

$$d_{new} = \lambda d, \quad (4)$$

i.e. a linear rescaling of all previous disparities.

2. IMAGE-DOMAIN-WARPING

Automatic view synthesis technology which synthesizes new views from 2-view video is highly popular. Such synthesis technology would be compatible to any S3D distribution infrastructure to the home. Image-domain Warping (IDW) is a view synthesis method which is able to automatically synthesize new views based on stereoscopic video input. In contrast to synthesis methods based on DIBR, which relies on dense disparity or depth maps, IDW employs only sparse disparities to synthesize a novel view. It use the facts that our human visual system is not able to very exactly estimate absolute depth and that it is not responsive to image distortions up to a certain level as long as images remain visually plausible, e.g. image distortions can be hidden in non-salient regions. They are used to compute an image warp which enforces desired sparse disparities in the final synthesized image while distortions are hidden in non-salient regions. To find out which disparities have to be enforced in a synthesized image.[Yin Zhao and Ce Zhu ,2011]

Goal: “warp” the pixels of the image so that they appear in the correct place for a new viewpoint. An advantage of IDW is, there is no need a geometric model of the object/environment can be done in time proportional to screen size and (Mostly) independent of object/environment complexity. Very less Disadvantage are their require Limited resolution and Excessive warping reveals several visual artifacts a linear rescaling of all previous disparities *d* is required to synthesize the new view.

In general, a stereo image pair doesn’t contain sufficient information to completely describe an image captured from a slightly different camera n position. IDW uses image saliency information to deal with this problem.[Aliaga ,2010]

I. Image Warps: We define a warp as a function that deforms the parameter domain of an image

$$W:[0, W] \times [0, H] \rightarrow R2 \quad (5)$$

Where *W* and *H* are the width and height of the image, respectively. Image warps have a long history of use in computer vision and graphics based problems goal of the IDW method is to compute a warping of Each of the initial stereo images that can be used to produce an output image meeting predefined properties (e.g. scaled image disparities). To do this, we formulate a quadratic energy functional *E (w)*. A warp w is then warps defined at regular grid positions.

$$W[p, q] := w(\Delta x\, p, \Delta y\, q). \quad (6)$$

E-ISSN: 2321–9637
Volume 2, Issue 1, January 2014
International Journal of Research in Advent Technology
Available Online at: http://www.ijrat.org

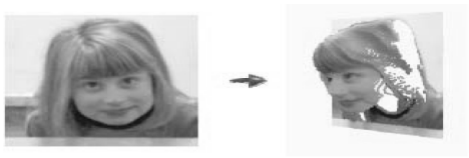

Fig.3 Example of a warping function that deforms an input image.

II. View Synthesis: The IDW algorithm computes N-view video from 2-view video where *N* >2. It can be separated into four modules, In the Warp Calculator only that warps are computed which are necessary for the Synthesis of A view that is located in the middle between two input views. These warps are calculated by minimizing their respective error functional. Calculated warps are then used in the Warp Interpolator/Extrapolator module to interpolate or extrapolate the warps that are necessary to synthesize the *N* output views, as required by a particular multi-view autosterescopic display. Finally, in the Image-domain warped module, images are warped to synthesize the output images.

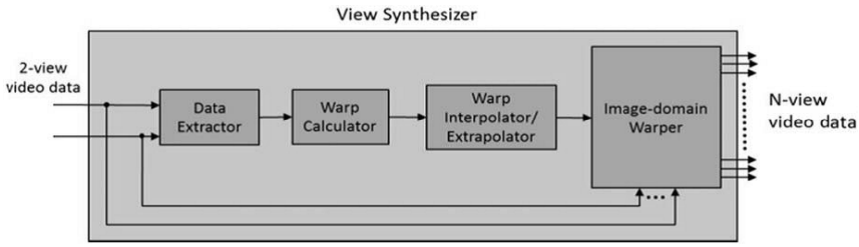

Fig. 4 Block diagram of the view synthesizer which converts 2-view videoto N-view video.

I. Data Extraction: First, a sparse set of disparity features is extracted. These sparse disparities are estimatedin an automatic, accurate and robust way.Disparities of vertical image edges are particularly important for the stereo sis, i.e. the perceived depth. For this reason, additional features and corresponding disparities are detected such that features lay uniformly distributed on nearly vertical image edges. Disparities of detected features are estimated using the Lucas-Kanade method. The availability of such features is also important to prevent salient synthesis errors with IDW like bending edges in the synthesized image

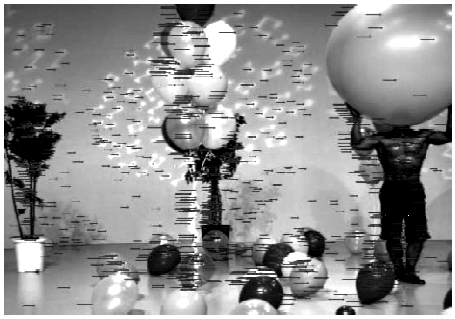

Fig. 5 Disparities estimated at sparse feature positions

II. Warp Calculation: Two warps *wL* and wR are computed which warp images *IL* and *IR*, respectively, to a camera position located in the center between the two input cameras. For a given sparse set of disparity features *(xL, xR)* ∈F, Disparities b

$$d = \frac{X_R - X_L}{2} \qquad (7)$$

E-ISSN: 2321–9637
Volume 2, Issue 1, January 2014
International Journal of Research in Advent Technology
Available Online at: http://www.ijrat.org

have to be enforced. Each warp is computed as the result of an energy minimization problem. Anenergy functional E is defined, and minimizing it yields a warp w that creates the desired change of disparities after view synthesis. The energy functional is defined with help of the extracted data and consists of three additive terms which are related to a particular type of constraint as described below. Each term is weighted with a parameter λ

$$E(W) := \lambda_d E_d(w) + \lambda_s E_s(w) + \lambda_t E_t(w) \quad (8)$$

In Warp calculation 4 constraints are used Disparity Constraints, Spatial Smoothness Constraints,Temporal Smoothness, Constraints and Energy Minimization.This functional always represents an over-constrained equation system. The number of spatial and temporal smoothness constraints is dense in the number of degrees of freedom of the warp Wto be solved, while the number of disparity constraints depends on the number of detected features, which is content dependent.

III. Warp Interpolation/Extrapolation: Multiview auto stereoscopic displays require many views from different camera positions as input. The main reason for this restriction is to reduce the overall computational complexity of the warp calculation. Furthermore, using this approach, the complexity of the warp computation does not depend on the number of output views required by a particular display system.

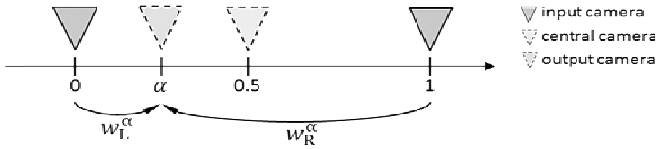

Fig.6 Cameras, camera positions, and associated warps

IV. Image-Domain Warping: An output image at a position α is synthesized according to i.e. $I\alpha$ is synthesized based on the input image which is closer to the desired output position. Because warps are continuous, no holes can occur in the synthesized image. In particular, open regions are implicitly unpainted by stretching unsealing texture from the neighborhood into the regionWe noticed that this kind of implicit unpointing provides good synthesis results in practice as long as views are synthesizedwhich are in the range $-0.5 \leq \alpha \leq 1.5$ (Fig.6). However, if only one image is used for the synthesis, empty regions can occur on the left or right border of the output image.

3. TRANSMISSION OF WARPS AS SUPPLEMENTARY DATA

To reduce the computational complexity at the receiver side, we modify the transmission system which was proposed in fig.7 The modified system is shown in fig.7 Thus, it is proposed to shift the warp extraction and warp calculation part to the sending side, and, in addition to the multiview data, to efficiently compress and transmit the warp calculation result, i.e. a restricted set of warps

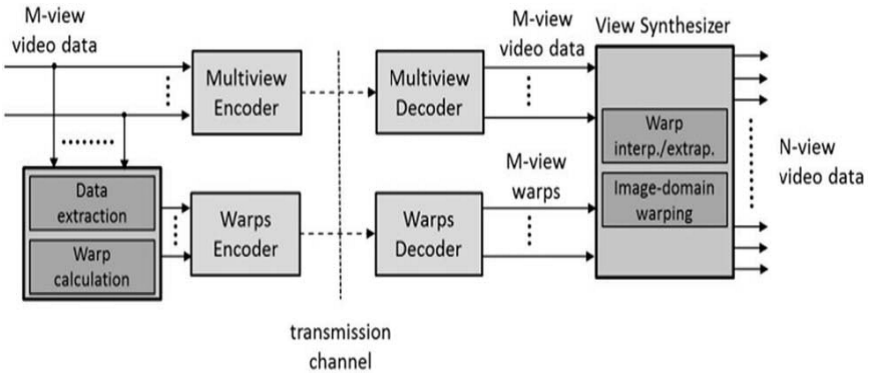

Fig.7 Modified transmission and view synthesis system

E-ISSN: 2321–9637
Volume 2, Issue 1, January 2014
International Journal of Research in Advent Technology
Available Online at: http://www.ijrat.org

I. Warp Coding with a Dedicated Warp Coder: Warps of all time instances and views are encoded successively. For each view, they are encoded separately and multiplexed into a single bit stream. Without loss of simplification, the coding of a warp sequence assigned to one view is described in the following. We denote the warp at time instant *f* as *w f* . Each warp *w f* is represented as a regular quad grid, in above (Fig.3) of fixed resolution where each node of the grid is indexed with integer coordinates *i, j* and has a 2D location *w f* [*i, j*] ∈R2 assigned

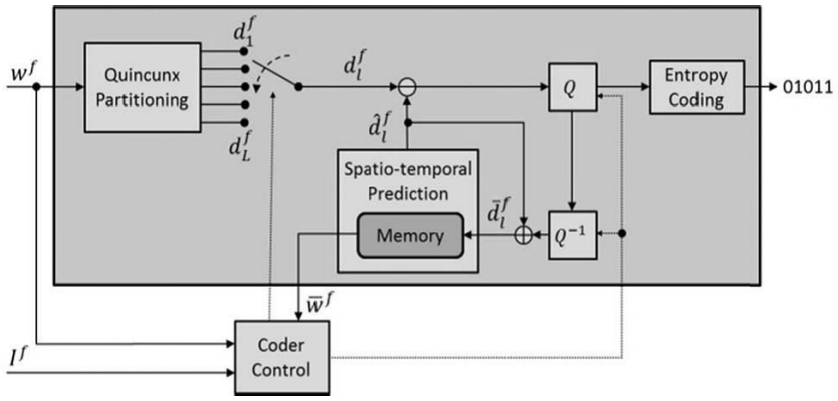

Fig. 8 Block-diagram of warp coder.

I. Spatial Partitioning: A partition consists of a set of 2D locations which we call a group of locations (GOLs). Each warp *w f* is partitioned into GOLs using a quincunx resolution pyramid,

II.Intra-Warp and Inter-Warp Prediction: Similar to video coding standards, three warp coding types and corresponding prediction modes are supported: INTRA, INTER_P and INTER_B. After prediction, residuals *w*[*i, j*] − ẃ*f* [*i, j*] are computed, uniformly quantized, and entropy coded. Quantized residuals of each GOL are entropy coded independently from other GOLs. In the INTRA prediction mode, all locations of *d f* 1 are scanned row-wise and predicted in a closed loop DPCM from previously scanned spatially neighbouring locations,as it is shown in Fig.9. Locations of all other GOLs *d fl* are predicted by the respective centroids computed from spatially neigh boring locations in $U_{k=1}^{l-1} d_l^f$ as indicated in Fig.9.

II. Warp Coding With Help of a Video Coder: To take advantage of already existing and highly sophisticated video coding technology, we propose the warp coding system show in Fig.10 as an alternative to the dedicated warp coding method shown.

I. Coding System: Similar to the coding with the dedicated warp coder, warps are encoded separately for each view and then multiplexed into a single bit stream. To encode a warp, first, a Warp Quantize is used to convert each warp into an 8-bit grayscale image representation.

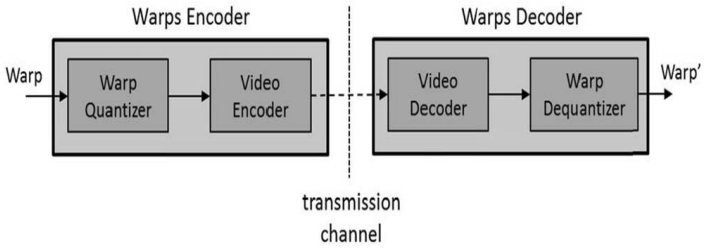

Fig.9. Warp coding system using a video coder.

E-ISSN: 2321–9637
Volume 2, Issue 1, January 2014
International Journal of Research in Advent Technology
Available Online at: http://www.ijrat.org

4. DISCUSSION

1) Recently, the Moving Pictures Experts Group (MPEG) issued a Call for Proposals (CfP) on 3D Video Coding technology with the goal to identify

- A 3D video format,
- A corresponding efficient compression technology,
- A view synthesis technology which enables an efficient synthesis of new views based on the 3D video format.

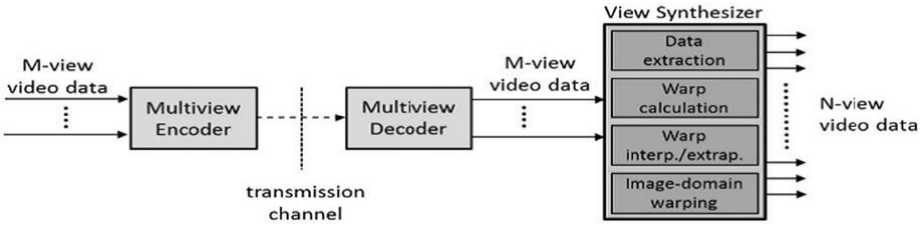

Fig. 10 Transmission and view synthesis system.

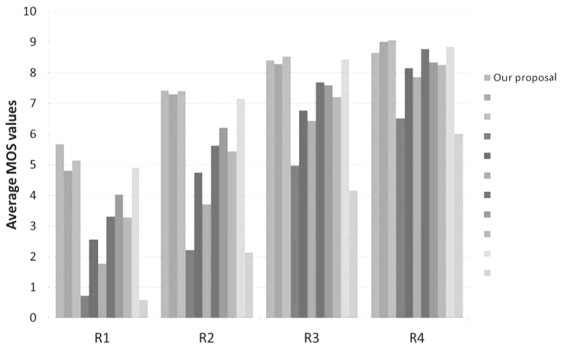

Fig..11. Assessed quality of the MPEG in the Multiview autosterescopic display test scenario.

This includes Data Extraction, Warp Calculation, and Warp Interpolation/Extrapolation. Please note that such a 3D format with 2 views is already supported by existing consumer and professional stereo cameras. With each proposal to the CfP, compressed bit streams, a decoder, and view synthesis software had to be provided. Bit streams had to be compressed at predefined target bit rates Proposals were evaluated by assessing the quality of the synthesized views through formal subjective testing on both stereoscopic and multiview auto stereoscopic displays. Fig.11 shows the quality assessed on a multiview auto stereoscopic display. Qualitatively similar results were also assessed on a stereoscopic display; the corresponding stereo sequences can be found here for download. [Sikora Thomas,1997]

5. CONCLUSION

A view synthesis method based on Image-domain-Warping. Its approach automatically synthesizes new views from stereoscopic 3D video. It relies on an automatic estimation of sparse disparities and image saliency information, and enforces target disparities in the synthesized images using an image warping framework. Image-domain Warping leads to high quality synthesis results without requiring depth map estimation and transmission. However a reuse of existing video coding technology has the advantage of reduced development and production costs. For this reason, JCT-3V planes to extend the upcoming 3D-HEVC standard by warp coding based on HEVC, which will allow receivers to perform a synthesis of new views based on Image-domain- Warping While a dedicated warp coder can have a stronger coding efficiency, the advantage of reusing video coding technology for warp coding lies the reduced development and production costs, i.e.in the reuse of available video coding chips. For this reason, based on the evaluation results presented in it, JCT-3V planes to extend the upcoming 3D-HEVC standard by warp coding based on HEVC. This will enable the transmission of

E-ISSN: 2321–9637
Volume 2, Issue 1, January 2014
International Journal of Research in Advent Technology
Available Online at: http://www.ijrat.org

multi-view video plus warp data with an internationalstandard, which will allow the use of IDW for view synthesisat the receiver side.

REFERENCES

[1] Aliaga G. Daniel image warping CS635 Spring 2010
[2] Fehn C, "Depth-image-based rendering (DIBR), compression, and Transmission for a new approach on 3D- TV," *Proc. SPIE*, vol. 5291, Vol. 93–104, May 2004
[3] Le Gall D.J, "The MPEG video compression algorithm," Signal Processing: Image Commune.1992, vol. 4, no. 4, pp. 129–140
[4] Muller K, Merkle P, and T. Wiegand, "3-D video representation using depth maps," *Proc. IEEE*, vol. 99, no. 4, pp. 643–656, Apr. 2011.
[5] Masayuki T, Zhao Y, and C. Zhu, *3D-TV* System with Depth- Image-Based Rendering: Architecture, Techniques and Challenges, 1st sed. New York, NY, USA: Springer-Verlag, 2012.
[6] Nikolce Stefanoski, Oliver Wang, Manuel Lang, Pierre Greisen, Simon Heinzle, and Aljosa Smolic IEEE Transactions On Image Processing, Vol. 22, No. 9, September 2013
[7] Smolic A,"3D video and free viewpoint video—From capture to display," Pattern Recognit., vol. 44, no. 9, pp. 1958–1968, Sep. 2011.
[8] Sikora Thomas, The MPEG-4 Video Standard Verification Model, Senior Member, IEEE Transactions On Circuits And Systems For Video Technology, Vol. 7, No. 1, February 1997
[9] Yin Zhao, Ce Zhu,Depth No-Synthesis-Error Model for View Synthesis in 3-D Video Senior Member, *IEEE*, Zhenzhong Chen, Member, *IEEE*, and Lu Yu, Member, *IEEE* IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 20, NO. 8, AUGUST 2011
[10] Zilly F, Riechert C, P. Eisert, and P. Kauff, "Semantic kernels binarized—A feature descriptor for fast and robust matching," in Nov. 2011, pp. 39–48.