V. Lakshmi Chetana, D. Prasad / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
Vol. 3, Issue 4, Jul-Aug 2013, pp.2189-2195

Color Space-Time Interest Points for Live Feed

V. Lakshmi Chetana*, D. Prasad**
*(Department of Computer Science & Engineering, DVR & Dr.HS MIC College of Technology, Kanchikacherla
** (Assoc. Prof. , Department of Computer Science & Engineering, DVR & Dr.HS MIC College of Technology, Kanchikacherla

ABSTRACT
Interest point detection is the vital aspect of computer vision and image processing. Interest point detection is very important for image retrieval and object categorization systems from which local descriptors are calculated for image matching schemes. Earlier approaches largely ignored color aspect as interest point calculation are largely based on luminance. However the use of color increases the distinctiveness of interest points. Subsequently an approach that uses saliency-based feature selection aided with a principle component analysis-based scale selection method was developed that happens to be a light-invariant interest points detection system. This paper introduces color interest points for sparse image representation. In the domain of video-indexing framework,interest points provides more useful information when compared to static images. Since human perception system is naturally influenced by motion of objects, we propose to extend the above approach for dynamic video streams using Space-Time Interest Points(STIP) method. The method includes an algorithm for scale adaptation of spatio-temporal interest points. STIP renders moving objects in a live feed and characterizes the specific changes in the movement of these objects. A practical implementation of the proposed system validates our claim to support live video feeds and further it can be used in domains such as Motion Tracking, Entity Detection and Naming applications that have abundance importance.

***Keywords* -** Interest points, Scale selection, image retrieval, STIP, LoG, Object Categorization

I. INTRODUCTION

Interest point detection is a vital research area in the field of computer vision and image processing. In particular image retrieval and object categorization heavily depend on interest point detection from which local image descriptors are calculated for image and object matching[1].

Majority of interest point extraction algorithms are purely based on luminance.These methods ignore the important information contained in the color channels.And hence interest point extraction based on luminance is not more effective.So color interest points are introduced.

Color plays a very important role in the pre-attentive stages of feature detection. Color provides extra information which allows the distinctiveness between various reasons of color variations, such as change due to shadows, light source reflections and object reflectance variations.[2] The detection and classification of local structures(i.e. edges, corners and T-junctions) in color images is important for many applications such as image segmentation, image matching, object recognition, visual tracking in the fields of image processing and computer vision[3],[4],[5]. Earlier approach of feature detectors uses the dense sampling representation of image in which the redundant information is also considered while extracting the features. In this paper we use a sparse representation of the image. The current trend in object recognition is increasing the number of points [6] by applying several detectors or by combining them [7][8] or making the interest point distribution as dense as possible[9].While such a dense sampling approach provides accurate object recognition, they basically change the task of discarding the non-discriminative points to the classifier.With the extreme growth of image and video data sets, clustering and offline training of features become less feasible[10]. By reducing the number of features and working with expected number of sparse features, larger image data sets can be processed in less time. Now the goal is to reduce the number of interest points extracted for obtaining better image retrieval or object recognition results.
Recent work has aimed to find unique features, i.e., by performing an evaluation of all features within the dataset or per image class and choosing the most frequent ones[11].For this, the best option is to use color to increase the distinctiveness of interest points[12][13]. This way may provide selective search for robust features reducing the total number of interest points used for image retrieval.

In this paper ,we propose color interest points to obtain a sparse image representation. With the sparse image representation we find the interest points which are sparsely scattered in the spatial domain. With this representation we can easily identify the object with minimum number of interest points.

To reduce the sensitivity to imaging conditions ,light-invariant interest points are proposed. To obtain light-invariant interest points , the quasi –invariant derivatives of the HSI color space are used. For color boosted points, color statistics are derived from the occurrence probability of colors. This way color boosted points are obtained through saliency based feature selection. Later a PCA-based scale selection method is proposed which gives interest points robust to scale.The use of color allows to extract repeatable and scale invariant interest points[21].

II. COMMON PIPELINE FOR IMAGE RETRIEVAL AND OBJECT CATEGORIZATION

Fig. 1 illustrates the common pipeline for image retrieval and object categorization

Feature extraction is carried out with either global or local features. In general, global features lack robustness against occlusions and cluttering and provide a fast and efficient way of image representation. Local features are either intensity- or color-based interest points. Earlier, dense sampling of local features has been used as it provides good performance, but the only problem is requires more number of interest points for feature extraction.[21]

Descriptors represent the local image information around the interest points. They can be categorized into three classes: They describe the distribution of certain local properties of the image [e.g., scale-invariant feature transform (SIFT)], spatial frequency (e.g., wavelets), or other differentials(e.g., local jets) [13] .For every feature extracted a descriptor is computed. A disadvantage is that the runtime increases with their number. Efficient ways to calculate these descriptors exist, e.g., for features with overlapping areas of support, previously calculated results can be used. [21]

Clustering for signature generation, feature generalization,or vocabulary estimation assigns the descriptors into a subset of categories. There are hierarchical and partitional approaches to clustering. Due to the excessive memory and runtime requirements of hierarchical clustering[14], partitional clustering such as k-means is the method of choice in creating feature signatures.

Matching summarizes the classification of images. Image descriptors are compared with previously learnt and stored models. This is computed by a similarity search or by building a model based on supervised or unsupervised learning techniques. Classification approaches need feature selection to discard irrelevant and redundant information [15]–[17]. It is shown that a powerful matching step can successfully discard irrelevant information, and better

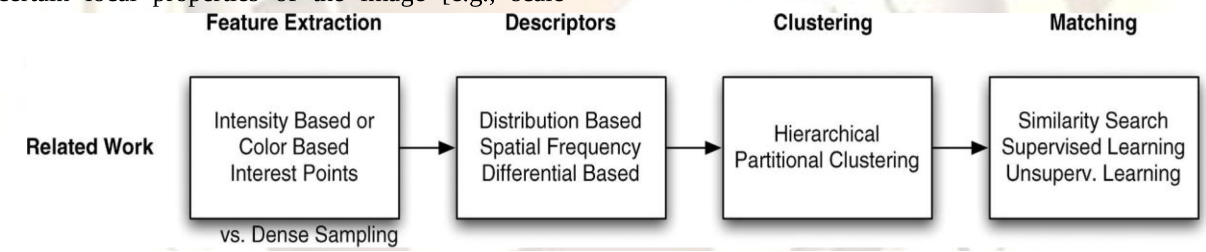

Fig. 1 Major steps for image retrieval and object categorizaton

performance is gained [9]. Training and clustering are the most time-consuming steps in state-of-the-art recognition frameworks.

Clustering of a global dictionary takes several days for current benchmark image databases, becoming less feasible for online databases resulting in several billions of features [10]. Therefore, one of our goals is a feature selection using color saliency within the first stage of this scheme. The use of color provides selective search reducing the total number of interest points used for image retrieval.

The important part in image retrieval and object categorization is feature extraction. So this paper mainly focuses on extracting the features in images. We extend the idea of interest points into spatio-temporal domain i.e., to the dynamic feed(video).The next sections describe about the color interest points both in spatial and spatio-temporal domain.

III. COLOR INTEREST POINTS IN SPATIAL DOMAIN

In this section, a scale invariant interest point detector is discussed: Harris-Laplace with color boosting. It is based on the Harris corner detector and use Laplacian scale selection [18].

3.1 *Interest points in spatial domain*

To extend the Harris corner detector to color information, the intensity-based version is taken. The adapted second moment matrix μ for position x is defined as follows [19]:

$$\mu(x,\sigma_I,\sigma_D) = \sigma_D^2 G(\sigma_I) \otimes \begin{pmatrix} L_x^2(x,\sigma_D) & L_xL_y(x,\sigma_D) \\ L_xL_y(x,\sigma_D) & L_y^2(x,\sigma_D) \end{pmatrix} \qquad (1)$$

Where $\otimes$ denotes the convolution

σ_I is the integration scale.
σ_D is the differentiation scale.
$G(\sigma_I)$ is the Gaussian kernel of size.
$L_x^2(x, \sigma_D)$ is the derivative computed in the x-direction at point x using using differentiation scale σ_D.
$L_y^2(x, \sigma_D)$ is the derivative computed in the y-direction at point x using differentiation scale σ_D.
$L_x L_y(x, \sigma_D)$ is the derivative computed in both the directions x & y at point x using the differentiation scale σ_D.

The matrix describes the gradient distribution in the local neighborhood of point x. Local derivatives are computed by Gaussian kernels with a differentiation scale denoted by σ_D . The derivatives are averaged in the neighborhood of point x by smoothing with a Gaussian window suitable for the integration scale σ_I.

The second moment matrix can be computed by a transformation0 in the RGB space[12].The first step is to determine the gradients of each component of the RGB color space. The gradients are then transformed into desired color space. By multiplying and summing of the transformed gradients, all components of the second moment matrix are computed. In symbolic form ,an arbitrary color space C is used with its n components $[c_{1},\ldots c_n]^T$ .The elements for μ are then calculated more generally as follows.[21]

$$L_x^2(x, \sigma_D) = \sum_{i=1}^{n} c_{i,x}^2(\sigma_D)$$

$$L_x L_y(x, \sigma_D) = \sum_{i=1}^{n} c_{i,x}(\sigma_D) c_{i,y}(\sigma_D)$$

$$L_y^2(x, \sigma_D) = \sum_{i=1}^{n} c_{i,y}^2(\sigma_D) \qquad (2)$$

Where $c_{i,x}$ and $c_{i,y}$ are the components of the transformed color channel gradients at scale σ_D .Subscripts x and y are the directions of the gradient.

The eigenvalues of the second moment matrix represent the two principal signal changes in the neighborhood of a point. Interest points are extracted where both eigenvalues are significant i.e. the signal change is significant in orthogonal directions, which is true for corners, junctions, etc.
The Harris corner detector [20] depends on the properties of the second moment matrix. It combines the trace and the determinant of the matrix into a cornerness measure:

$$cornerness(H) = \det(\mu(x, \sigma_I, \sigma_D)) - K trace^2(\mu(x, \sigma_I, \sigma_D)) \qquad (3)$$

where K is an empirical constant with values between 0:04 and 0:06. Local maxima of the cornerness measure (equation 3) determine the interest point locations.

This leads to stable locations that are robust to noise and scale changes upto a factor of 1.414,translation,rotation under different color spaces .For many computer vision systems it is crucial to find the scale-invariant points.So in the next section ,an approach for scale selection for local features in different color spaces is proposed.

3.2 *Scale selection in spatial domain*

Using the elements of second moment matrix, we extend the color Harris to the Harris-Laplacian by applying Harris corners to LoG at different scales. Automatic scale selection allows for the selection of the characteristic scale of a point, which depends on the local structure neighborhood point. The characteristic scale is the scale for which a given function attains a maximum over scales. It has been shown that the cornerness measure of the Harris corner detector rarely attains a maximum over scales. Thus, it is not suitable for selecting a proper scale. However, the Laplacian-of-Gaussian(LoG) attain maximum over scales. With σ_n, the scale parameter of the LoG, it is defined for a point x as:[24]

$$|LoG(x, \sigma_n)| = \sigma_n^2 \left| L_{xx}(x, \sigma_n) + L_{yy}(x, \sigma_n) \right| \qquad (4)$$

The function reaches a maximum when the size of the kernel matches the size of the local structure around the point.

IV. COLOR INTEREST POINTS IN SPATIO-TEMPORAL DOMAIN

4.1 *Interest points in spatio-temporal domain*

In this section, we develop an operator that responds to events in temporal image sequences at specific locations and with specific extents in space-time. The idea is to extend the concept of interest points in the spatial domain by requiring the image values in local spatio-temporal domains to have large variations along both the spatial and temporal directions. Points with such properties are the spatial interest points with different locations in time corresponding to a local spatio-temporal neighbourhoods with non-constant movement.

To model a spatio-temporal image sequence ,we use a function f:$R^2 \times R \rightarrow R$ and construct its linear scale-space representation L : $R^2 \times R \times R^2_+ \rightarrow R$ by convulting f with Gaussian kernel with different spatial variance σ_l^2 and temporal variance τ_l^2

$$L(.; \sigma_l^2, \tau_l^2) = G(.; \sigma_l^2, \tau_l^2) * f(.) \qquad (5)$$

Where the spatio-temporal seperable Gaussian kernel is defined as

$$G(x, y, t; \sigma_l^2, \tau_l^2) = \frac{1}{\sqrt{(2\pi)^3 \sigma_l^4 \tau_l^2}} \times e^{\left(\frac{-(x^2+y^2)}{2\sigma_l^2} - \frac{t^2}{2\tau_l^2}\right)} \qquad (6)$$

Separate scale parameter for the temporal domain is essential because the spatial and the temporal scope of events are independent.

Similar to the spatial domain, we consider a spatio-temporal second moment matrix,which is a 3×3 matrix composed of first order spatial and temporal derivatives averaged using a Gaussian kernel $G(.;\sigma_l^2,\tau_l^2)$ [22]

$$\mu = G(.;\sigma_i^2,\tau_i^2) * \begin{pmatrix} L_x^2 & L_xL_y & L_xL_t \\ L_xL_y & L_y^2 & L_yL_t \\ L_xL_t & L_yL_t & L_t^2 \end{pmatrix} \quad (7)$$

Where the integration scales σ_i^2 and τ_i^2 are related to local scales by $\sigma_i^2 =_{S.} \sigma_l^2$ and τ_i^2=s. τ_l^2 while the first order derivatives are defined as[23]

$$L_\epsilon(.;\sigma_l^2,\tau_l^2) = \partial_\epsilon(G * f) \quad (8)$$

To find interest points,we search for regions in f having significant eigenvalues λ_1, λ_2 and λ_3 of the second moment matrix for spatio-temporal domain µ. The Harris cornerness measure in spatio-temporal domain is defined as follows

$$cornerness(H) = \det(\mu) - K trace^3(\mu) \quad (9)$$

The following fig. 2 . illustrates detection of interest points in spatial domain for a moving corner.

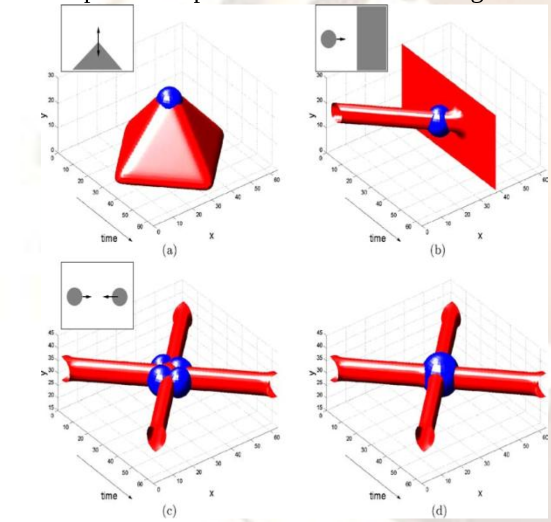

Fig.2. Results of detecting spatio-temporal interest

points on synthetic image sequences: (a) Moving corner;(b) A merge of a ball and a wall; (c) Collision of two balls with interest points detected at scales σ_l^2 = 8 and τ_l^2 = 8;(d) the same as in (c) but with interest points detected at scales σ_l^2 = 16 and τ_l^2 = 16

4.2 *Scale Selection in Spatial-Temporal Domain*

To estimate the spatio- temporal extent of an event in space-time, we follow works on local scale selection proposed in the spatial domain by Lindeberg [26] as well as in the temporal domain [25]. For the purpose of analysis, we will first study a prototype event represented by a spatio-temporal Gaussian blob

$$f(x,y,t;\sigma_0^2,\tau_0^2) = \frac{1}{\sqrt{(2\pi)^3\sigma_l^4\tau_l^2}} \times e^{\left(\frac{-(x^2+y^2)}{2\sigma_0^2} - \frac{t^2}{2\tau_0^2}\right)} \quad (10)$$

with spatial variance σ_0^2 and temporal variance τ_0^2(see figure).Using the semi-group property of the Gaussian kernel,it follows that the scale-space representation of ***f*** is

$$L(x,y,t,\sigma^2,\tau^2) = G(x,y,t,\sigma_0^2+\sigma^2,\tau_0^2+\tau^2) \quad (11)$$

To recover the spatio-temporal extent (σ_0, τ_0) of *f* we consider the scale-normalized spatio-temporal Laplacian operator defined by

$$\nabla^2_{norm}L = L_{xx,norm} + L_{yy,norm} + L_{tt,norm} \quad (12)$$

where $L_{xx,norm} = \sigma^{2a}\tau^{2b}L_{xx}$

$L_{yy,norm} = \sigma^{2a}\tau^{2b}L_{yy}$

$L_{tt,norm} = \sigma^{2a}\tau^{2b}L_{tt}$

As shown in [23], given the appropriate normalization parameters *a* = 1, *b* = 1/4, *c* = 1/2 and *d* = 3/4, the size of the blob *f* can be estimated from the scale values $\tilde{\sigma}^2$ and $\tilde{\tau}^2$ for which $\nabla^2_{norm}L$ assumes local extrema over scales, space and time. Hence, the spatio-temporal extent of the blob can be estimated by detecting local extrema of

$$\nabla^2_{norm}L = \sigma^2\tau^{1/2}(L_{xx} + L_{yy}) + \sigma\tau^{3/2} L_{tt}. \quad (13)$$

Over both spatial and temporal scales.

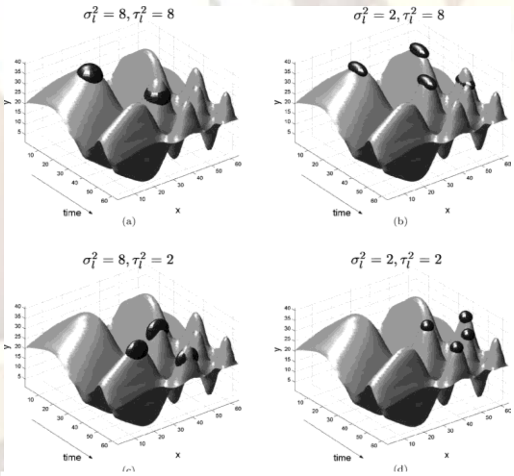

Fig. 3. Results of detecting interest point at different spatial and temporal scales for a synthetic sequence with impulses having varying extents in space and time. The extents of the detected events roughly corresponds to the scale parameters σ^2 and τ^2 used for computing *H* (9).

4.3 *Scale-adapted space-time interest points*

Local scale estimation using the normalized Laplace operator has shown to be very useful in the spatial domain [26, 27]. In particularly, Mikolajczyk and Schmid [28] combined the Harris interest point operator with the normalized Laplace operator and derived the scale-invariant Harris-Laplace interest point detector. The idea is to find points in scale-space that are both maxima of the Harris function corner function H (4) in space and extrema of the scale-normalized spatial Laplace opera-tor over scale.

Here, we extend this idea and detect interest points that are simultaneous maxima of the spatio-temporal corner function H (9) as well as extrema of the normalized spatio-temporal Laplace operator $\nabla^2_{norm}L$ (13). Hence, we

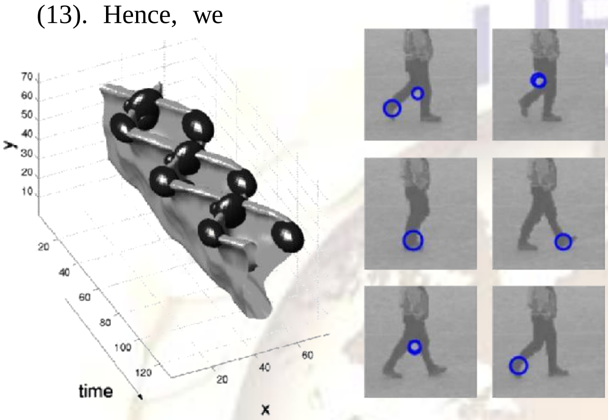

Fig. 4. Results of detecting spatio-temporal interest points for the motion of the legs of a walking person: (a) 3-D plot with a threshold surface of a leg pattern (up side down) and detected interest points; (b) interest points over-laid on single frames in the sequence.

detect interest points for a set of sparsely distributed scale values and then track these

points in spatio -temporal scale-time-space towards the extrema of $\nabla^2_{norm}L$. We do this by iteratively updating the scale and the position of the interest points by

(i) Selecting the neighboring spatio-temporal scale that maximizes $(\nabla^2_{norm}L)^2$ and

(ii) Re-detecting the space-time location of the interest point at a new scale until the position and the scale converge to the stable values [23].

To illustrate the performance of the scale-adapted spatio-temporal interest point detector, let us consider a sequence with a walking person and non-constant image velocities due to the oscillating motion of the legs. As can be seen in fig. 3, the pattern gives rise to stable interest points. Note that the detected points are well-localized both in space and time and correspond to events such as the stopping and starting feet. From the space-time plot in fig. 3(a), we can also observe how the selected spatial and temporal scales of the detected features roughly match the spatio-temporal extents of the corresponding image structures [29].

V. RESULTS

Fig. 5. The above row represents identification of interest points based on color. The below row represents identification of interest points without color(i.e based on luminance)

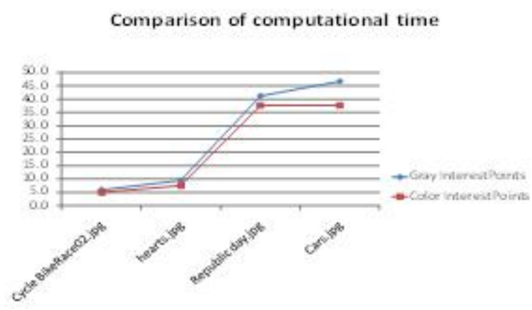

Fig. 6 .Comparison of computational times for color interest points and gray-level interest points for different types of

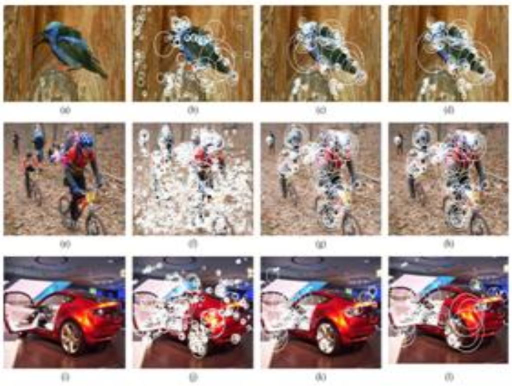

Fig. 7. Three visual examples of the VOC 2007 data set (rows 1–3). The original images are found in the first column, luminance-based Harris–Laplacian in the second, light-invariant points in the third, and color boosted points in the fourth. White circles indicate the location and the scale of interest points; parameters are chosen equal to the ones used in the experiments with 500 as the suggested threshold for Harris–Laplacian and a maximum of 400 interest points for the proposed approaches. (a) Bird. (b) Harris–Laplacian.(c)Light-invariantpoints.(d)Colorboostedoints. (e) Cyclic Bike race 02. (f) Harris–Laplacian. (g) Light-invariant points. (h) Color boosted points. (i) Cars(3) (j) Harris–Laplacian. (k) Light-invariant points. (l) Color boosted points.

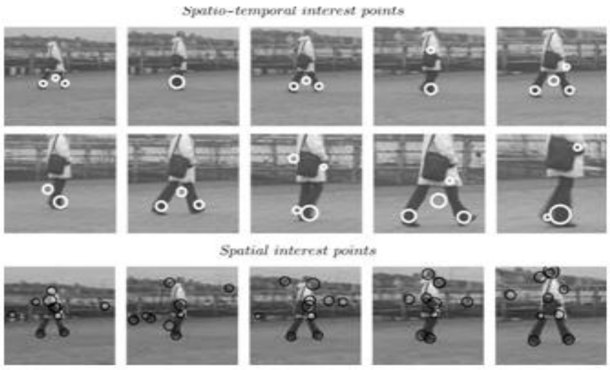

Fig 8. Top: Results of spatio-temporal interest point detection for a zoom-in sequence of a walking person. The spatial scale of the detected points (corresponding to the size of circles) matches the increasing spatial extent of the image structures and verifies the invariance of the interest points with respect to changes in spatial scale. Bottom: Pure spatial interest point detector (here, Harris-Laplace. Mikolajczyk and Schmid 2001) selects both moving and stationary points in the image sequence

VI . CONCLUSION

In this paper, an approach has been proposed to extract the light invariant interest points based on color for static images. we have extended this interest points to dynamic content that is for videos which play a very important role for Object or Scene identification and retrieval and Motion Capturing. We have described an interest point detector that finds local image features in space –time characterized by a high variation of the image values in space and varying motion over time. The localization of the interest points of each frame is identified by the Harris Detector and Laplacian operator is applied to these points to find scale-invariant points in space-time. The combination of these two operators is combined as Space –Time Interest Points.

In future work, this application can be extended to the field of motion-based recognition and event detection. Generally the present motion capturing systems(Eg :CC cameras used in airports, shopping malls etc..;) are high in luminance content .But in gray level images or videos most of the important content is discarded due to their less saliency. So usage of this color STIP's can improve the detection or recognition or identification of objects.

REFERENCES

[1] R. Fergus, P. Perona, and A. Zisserman, "Object class recognition by unsupervised scale-invariant learning," in *Proc. CVPR*, 2003, pp. II-264–II-271.

[2] Joost van de Weijer,"Color Features and Local Structure in Images".

[3] Cordelia Schmid, Roger Mohr and Christian Bauckhage ,"Evaluating of interest points detectors", published in "International Journal of Computer Vision 37, 2 (2000) 151--172"

[4] Jianbo Shi and Carlo Tomasi,"Good features to track",research report submiited in December 2005.

[5] SILVANO DI ZENZO ,"A note on the gradient of a multi-image", COMPUTER VISION, GRAPHICS, AND IMAGE PROCESSING 33, 116-125 (1986)

[6] J. Zhang, M. Marszałek, S. Lazebnik, and C. Schmid, "Local features and kernels for classification of texture and object categories: A com-prehensive study," *Int. J. Comput. Vis.*, vol. 73, no. 2, pp. 213–238, Jun. 2007.

[7] K. Mikolajczyk, B. Leibe, and B. Schiele, "Multiple object class de-tection with a generative model," in *Proc. CVPR*, 2006, pp. 26–36.

[8] J. Sivic, B. Russell, A. A. Efros, A. Zisserman, and B. Freeman, "Dis-covering objects and their location in images," in *Proc. ICCV*, 2005, pp. 370–377.

[9] T. Tuytelaars and C. Schmid, "Vector quantizing feature space with a regular lattice," in *Proc. ICCV*, 2007, pp. 1–8.

[10] G. Schindler, M. Brown, and R. Szeliski, "City-scale location recognition," in *Proc. CVPR*, 2007, pp. 1–7.

[11] P. Turcot and D. G. Lowe, "Better matching with fewer features: The selection of useful features in large

database recognition problems," in *Proc. ICCV Workshop*, 2009, pp. 2109–2116.

[12] J. van de Weijer and T. Gevers, "Edge and corner detection by photo metric quasi-invariants," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 4, pp. 625–630, Apr. 2005.

[13] Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005.

[14] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," *ACM Comput. Surv.*, vol. 31, no. 3, pp. 264–323, Sep. 1999.

[15] R. Okada and S. Soatto, "Relevant feature selection for human pose estimation and localization in cluttered images," in *Proc. ECCV*, 2008, pp. 434–445.

[16] G. Dorko and C. Schmid, "Selection of scale-invariant parts for object class recognition," in *Proc. ICCV*, 2003, pp. 634–639.

[17] F. Jurie and B. Triggs, "Creating efficient codebooks for visual recognition," in *Proc. ICCV*, 2005, pp. 604–610.

[18] Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. V. Gool. A comparison of affine region detectors. International Journal of Computer Vision, 65(1-2):43–72, 2005.

[19] K. Mikolajczyk and C. Schmid. Scale and affine invariant interest point detectors. International Journal of Computer Vision, 60(1):63–86, 2004.

[20] C. Harris and M. Stephens. A combined corner and edge detector. In Proceedings of The Fourth Alvey Vision Conference, Manchester, pages 147–151, 1988.

[21] Julian Stöttinger, Allan Hanbury, Nicu Sebe, "Sparse color interest points for image retrieval and object categorization", IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 21, NO. 5, MAY 2012

[22] Ivan Laptev,"On Space-Time Interest Points" in International Journal of Computer Vision 64(2/3),107-123,2005.

[23] I. Laptev and T. Lindeberg. Interest point detection and scale selec-tion in space-time. In L.D. Griffin and M Lillholm, editors, *Scale-Space'03*, volume 2695 of *Lecture Notes in Computer Science*, pages 372–387. Springer Verlag, Berlin, 2003.

[24] Koen E.A. van de Sande, Theo Gevers and Cees G.M. Snoek "Color descriptors for object categorization" IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 32 (9), page 1582--1596, 2010.

[25] T. Lindeberg. On automatic selection of temporal scales in time-causal scale-space. In *AFPAC'97: Algebraic Frames for the Perception-Action Cycle*, volume 1315 of *Lecture Notes in Computer Science*, pages 94–113. Springer Verlag, Berlin, 1997.

[26] T. Lindeberg."Feature detection with automatic scale selection". *IJCV*, 30(2):77–116, 1998.

[27] O. Chomat, V.C. de Verdiere, D. Hall, and J.L. Crowley. Local scale selection for Gaussian based description techniques. In *Proc. ECCV*, volume 1842 of *Lecture Notes in Computer Science*, pages I:117–133, Dublin, Ireland, 2000. Springer Verlag, Berlin.

[28] K. Mikolajczyk and C. Schmid. Indexing based on scale invariant interest points. In *Proc. ICCV*, pages I:525–531, Vancouver, Canada, 2001.

[29] Ivan Laptev and Tony Lindeberg," Space-time Interest Points",in Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV'03).