Multi-dimensional Scale Saliency Feature Extraction Based on Entropic Graphs

P. Suau and F. Escolano

Robot Vision Group, Departamento de Ciencia de la Computación e IA
Universidad de Alicante, Spain
{pablo,sco}@dccia.ua.es

Abstract. In this paper we present a multi-dimensional version of the Kadir and Brady scale saliency feature extractor, based on Entropic Graphs and Rényi alpha-entropy estimation. The original Kadir and Brady algorithm is conditioned by the curse of dimensionality when estimating entropy from multi-dimensional data like RGB intensity values. Our approach naturally allows to increase dimensionality, being its computation time slightly affected by the number of dimensions. Our computation time experiments, based on hyperspectral images composed of 31 bands, demonstrate that our approach can be applied to computer vision fields, i.e. hyperspectral or satellite imaging, that can not be solved by means of the original algorithm.

1 Introduction

Low-level vision is a key step in many vision tasks whose results strongly depend on the quality of the information gathered from early stages. Feature extraction in computer vision is still a challenging problem, with open issues like generalization (the application of feature extraction to more complex data than grayscale intensities) or development of affine invariant detectors [1]. Several approaches have been presented in order to achieve some generalization in saliency detection, commonly based on statistical learning from multiple examples. Park *et al.* [2] build saliency maps from color and orientation features using Independent Component Analysis, in order to extract salient paths. In [3], color distinctiveness is also incorporated into saliency detection; from the analysis of color statistics, a color boosting method is proposed to incorporate color information to multiple saliency methods. More recently [4] combines information from the color channels for selection of scale invariant keypoints.

Our work is focused in Kadir and Brady scale saliency algorithm [5]. Although recent surveys [1][6] criticize this algorithm in performance compared to other methods, it is commonly applied to problems like robot localization [7] and object recognition [8]. Kadir and Brady pointed out that, unlike other feature extractors, their algorithm is able to calculate saliency from a set of different measures from the image, and not only grayscale intensity [5]. Due to the fact that probability density function (pdf) estimation, from which saliency is calculated, is based on histograms, increasing dimensionality is as simple as

G. Bebis et al. (Eds.): ISVC 2008, Part II, LNCS 5359, pp. 170–180, 2008.

using higher dimensional histograms. As a consequence, no learning is needed; however, such method is highly conditioned by the curse of dimensionality. We propose a new method for scale saliency feature extraction, by means of alternative estimators of saliency and divergence based on entropic graphs [9][10][11]. Complexity of these estimators is slightly affected by data dimensionality.

Although the claimed decrease of complexity order of our approach is not theoretically proven, our experiments with 31 band hyperspectral images clearly demonstrate this fact. The main difference with respect to Kadir and Brady color based extensions is that our approach may naturally handle higher dimensionalities. Our method is not efficient for gray level or colour images; however, it may be applied to problems, like hyperspectral or satellite images analysis, that can not be solved using original Kadir and Brady algorithm or any of the existing colour generalizations. This paper is structured as follows. In section 2, the original Kadir and Brady algorithm is summarized. Sections 3 and 4 describe how entropic graphs may be used to estimate Shannon entropy and divergence, the two main processes in scale saliency algorithm. Section 5 describes our approach, dealing in section 6 with some implementation considerations. In section 7, experimental results are shown. Finally, conclusions are presented in section 8.

2 The Kadir and Brady Scale Saliency Feature Extraction

Visual saliency may be defined as a measure of local complexity or unpredictability [5]. Salient features are distinctive, due to this local unpredictability, and have proved to be useful in the context of image registration. Using Shannon entropy, Gilles formulated local saliency in terms of local intensity histograms [12]. Given a pixel x, a local neighborhood R_x, and a descriptor D that takes values $\{d_1, ..., d_r\}$ (e.g. in an 8 bit grey level image D would range from 0 to 255), local entropy is defined as:

$$H_{D,R_x} = -\sum_i P_{D,R_x}(d_i) \log_2 P_{D,R_x}(d_i) \ . \tag{1}$$

where $P_{D,R_x}(d_i)$ is the probability of descriptor D taking the value d_i in the local region R_x.

However, this approach may be improved in many ways. The main drawback is that scale ($|R_x|$ in the latter equation) is a pre-selected parameter, so this model is only proper for images that contain features existing over a small range of scales. In order to solve this problem and others, Kadir and Brady proposed their scale saliency algorithm [5], extending saliency to work through scale space as well as through feature space; their approach is based on detecting salient features that exist over a narrow range of scales. This method can be summarized as follows: for each pixel x, local entropy H_D (Eq. 2) is calculated for each scale s between s_{min} and s_{max}; the scales S_p (Eq. 3) at which the entropy is a local maximum (is peaked) are chosen, and then the entropy is weighted (W_D, Eq. 4) at such scales by some measure of the self-dissimilarity in scale-space of the feature. The algorithm yields a sparse three dimensional array of scalar values Y_D (Eq. 5), containing weighted local entropies for all pixels at those scales where entropy is peaked.

$$H_D(s,x) = -\sum_{d\in D} P_{d,s,x}\log_2 P_{d,s,x}\ . \tag{2}$$

$$S_p = \{s : H_D(s-1,x) < H_D(s,x) > H_D(s+1,x)\}\ . \tag{3}$$

$$W_D(s,x) = \frac{s^2}{2s-1}\sum_{d\in D}|P_{d,s,x} - P_{d,s-1,x}|\ . \tag{4}$$

$$Y_D(s_p,x) = H_D(s_p,x)W_D(s_p,x)\ . \tag{5}$$

Kadir presented some generalizations of his scale saliency algorithm, including an isotropic requirement relaxation [13] and a color saliency model [14]. Both extensions of the original algorithm produce an increase of the time complexity order and memory allocation requirements. For instance, color saliency is estimated increasing the dimensionality of the intensity histograms from which the entropy is calculated (two dimensional histograms for the YC_bC_r color model, and three dimensional histograms for the RGB color model), exponentially increasing computation time and allocated memory. If even more information is added, the problem may be computationally prohibitive.

3 Shannon Entropy Estimation from Entropic Graphs

Entropic graphs are an alternative to common entropy estimation methods based on histograms. Although the equations introduced in this paper were first applied to estimate entropy and dissimilarity from Minimal Spanning Trees (MSTs) [11], our final implementation relies on KNNGs (K-Nearest Neighbors Graphs) to compute these estimations. This choice, that will be justified in section 6, is supported by the work of Neemuchwala et al. [10]. Let G be a graph formed by a set of vertices or nodes $X_n = \{x_1,\ldots,x_n\}$ with $x_i \in \mathbb{R}^d$ and edges $\{e\}$ that connect vertices in graph, being $e_{ij} = (x_i, x_j)$. If we denote by $M(X_n)$ the possible sets of edges in the class of graphs built from X_n where each vertex is connected to other K vertices in the graph (K neighbor graph), total edge length functional of the Euclidean power weighted K-Nearest Neighbor graph is

$$L_\gamma^{KNNG}(X_n) = \min_{M(X_n)}\sum_{e\in M(X_n)}|e|^\gamma\ . \tag{6}$$

with $\gamma \in (0,d)$ and $|e|$ the euclidean distance between connected graph vertices. It is intuitive that the length of the KNNG spanning a more concentrated nonuniform set of samples increases at a slower rate than does the KNNG spanning a uniformly distributed set of samples. This fact motivates the application of the KNNG as a way to test for randomness of data.

In [9], a method to approximate Shannon entropy from a MST is proposed, based on Rényi α-entropy. Our preliminary experiments show that it can be applied to KNNGs. In this case, the estimation is given by:

$$H_\alpha(X_n) = \frac{d}{\gamma}\left[ln\frac{L_\gamma(X_n)}{|e|^\alpha} - ln\beta_{L_\gamma,d}\right] . \quad (7)$$

that is an asymptotically unbiased and almost surely consistent estimator of the α-entropy of X_n, being $ne = |e_{ij}|, \alpha = (d-\gamma)/\gamma$ and $\beta_{L_\gamma,d}$ a constant bias criterion depending on the graph minimization criterion, but independent of X_n. We can estimate $H_\alpha(X_n)$ for different values of α by changing the edge weight exponent γ and recalculating $L_\gamma(X_n)$, without recalculating the KNNG from X_n.

The actual estimation of Shannon Entropy from $H_\alpha(X_n)$ can be achieved when $\alpha = 1$. However, α-entropy estimation from entropic graphs is only suitable for $\alpha \in [0, 1[$. In [9], an α^* approximation is given in order to search for the assymtotically $H_\alpha(X_n)$ value when $\alpha = 1$. This approximation was obtained empirically and is given by

$$\alpha^* = 1 - \frac{1.271 + 1.3912 \cdot exp^{-0.2499 \cdot d}}{n} . \quad (8)$$

The only preprocessing step needed is to calculate a α^* value for each scale, in the range between s_{min} and s_{max}, from Eq. 8. These values are used to estimate Shannon entropy for each multivalued pixel at each scale on an image. Our experiments show that this estimator was appropriate for including it in scale saliency computation. Furthermore, an important fact is that its complexity is slightly affected by data dimensionality.

4 Dissimilarity Estimation from Entropic Graphs

During Kadir scale saliency algorithm, entropy peaks in scale space must be weighted by means of a measure of dissimilarity between scales. In [15], Friedman and Rafsky's estimation of the Henze-Penrose divergence is presented as a method to approximate this dissimilarity by means of MSTs. As in the case of entropy estimation introduced in Section 3, the complexity of Friedman and Rafsky's method is slightly affected by data dimensionality.

Henze-Penrose divergence [16] between distributions f and g is defined as

$$D_{HP}(f||g) = \int \frac{p^2 f^2(x) + q^2 g^2(x)}{pf(x) + qg(x)} dx . \quad (9)$$

where the weights p and $q = 1 - p$ are selected in the interval $(0, 1)$. In order to estimate this divergence from two sets of multi-dimensional nodes $X = \{x_i\}$ and $O = \{o_i\}$, Friedman-Rafsky test [17] may be applied (an example is shown in Fig. 1):

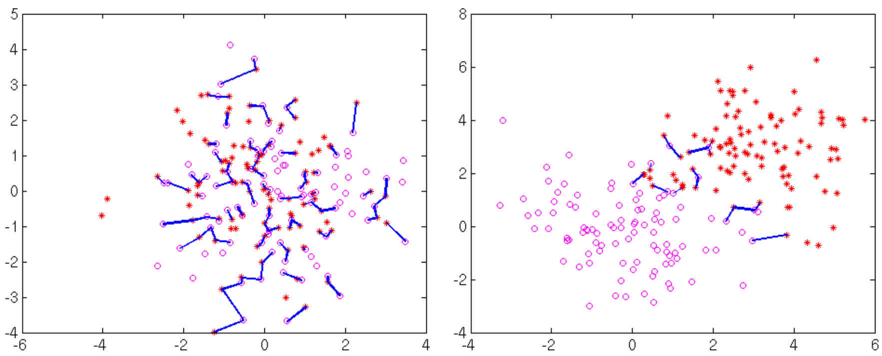

Fig. 1. Examples of Friedman and Rafsky's estimation of Henze-Penrose divergence for the case of two Gaussian densities. Left: the two densities have the same mean and variance parameters ($D_{HP}(O||X) = 0.485$). Right: the two densities have different mean values ($D_{HP}(O||X) = 0.925$).

- Build the MST from $\{x_i\} \bigcup \{o_i\}$.
- Retain only those edges that connect an $\{x_i\} \in X$ vertex to an $\{o_i\} \in O$ vertex.
- The test value is defined as the number of edges retained, *normalized* by the total number of edges of the MST built in the first step. This value is a consistent estimate of $1 - D_{HP}(O||X)$.

After checking in our preliminary experiments that this test may also be applied in the case of KNNGs, incorporating it in our scale saliency approach is straightforward. Let s_i the scale where an entropy peak has been found in scale space for a pixel. In order to weight this entropy value, dissimilarity with s_{i-1} must be calculated [5]. Let $M(X_{m,i})$ and $M(X_{n,i-1})$ be the KNNGs used to estimate entropy at scales s_i and s_{i-1} (being $m > n$ in general); it is not necessary to calculate a new KNNG to estimate dissimilarity between scales s_i and s_{i-1} since $X_{n,i-1} \subset X_{m,i}$ (new feature vectors are added to previous ones when increasing scale). Therefore, $M(X_{m,i})$ is the KNNG of $X_{n,i-1} \bigcup X_{m,i}$. Thus, the only calculation needed is to count the number of edges of $X_{m,i}$ that connect a node from $X_{m,i}/X_{n,i-1}$ to a node of $X_{n,i-1}$.

5 Scale Saliency from Entropic Graphs

Our approach for high-dimensional scale saliency from entropic graphs is summarized in Algorithm 1. In order to generalize original scale saliency algorithm [5] for multi-dimensional features apart from grayscale intensity, it uses entropic graphs to estimate saliency and divergence between scales. Before applying the algorithm, d features are calculated for each pixel on the image. These features could be color intensity, orientation and gradient magnitude, or any set of features we are interested to extract salience from. For each pixel at each scale a

KNNG is calculated (steps 1 and 2 of the algorithm). These KNNGs will be used to estimate saliency by means of the method explained in Section 3 (step 3). When an entropy peak is found in any scale, Henze-Penrose divergence is estimated using Friedman and Rafsky's method from the KNNG at that scale, as stated in Section 4 (Step 4). Entropy peak is then weighted by this dissimilarity measure.

```
Input: A d-dimensional array I containing d features for each pixel of the image
Output: An array HW containing weighted entropy values for all pixels on
        image at each scale
foreach pixel j of image do
    foreach scale s_i between s_min and s_max do
        (1) Create d-dimensional nodes X_i from the local neighborhood of pixel j
        at scale s_i in I;
        (2) Calculate L_γ(s_i)(X_i);
        (3) H(s_i) = H_α*(s_i)(X_n);
        if i > s_min + 1 then
            if H(s_{i-2}) < H(s_{i-1}) > H(s_i) then
                (* Entropy peak *)
                (4) W = D_HP(X_{i-2}||X_{i-1});
                (5) HW(s_{i-1}, j) = H(s_{i-1}) · W;
            end
            else
                (6) HW(s_{i-1}, j) = 0;
            end
        end
    end
end
```

Algorithm 1. Multi-dimensional scale saliency from entropic graphs

6 Implementation Considerations

As can be seen in Algorithm 1, the temporal efficiency of our approach strongly depends on the algorithm used to build an entropic graph from the data, due to the fact that an entropic graph must be computed for each pixel at each scale. An usual choice is to perform these estimations from MSTs [11]. However, two main issues arise when MSTs are used: the lack of practical MST update algorithms and the need of adding random noise to edge lengths in order to avoid unstable entropy estimation [10]. These two limitations would strongly affect computation time and the quality of the multi-dimensional feature extractor, and we overcome them choosing a different family of entropic graphs for our implementation: KNNGs.

If feature vectors extracted from image pixels are quantized, the probability of zero valued edges highly increases, yielding low length MSTs that may result in negative alpha-entropy values (Eq. 7) and unstable entropy estimation [10]. A proposed remedy is to add a small amount of uniform noise to the pixel feature values. However, during our experiments with KNNGs we noticed that if the value of K, and as a consequence the number of edges of the graph, is high enough, the probability of obtaining negative entropy values tends to zero. Thus, a clear advantage of entropy estimation from KNNGs is its increased accuracy compared to MSTs. We assigned to the parameter K the value 3 during all our experiments.

Regarding computation time, it can be remarkably decreased if during the scale-space entropy computation for any pixel, the implementation avoids to reconsider points of its neigborhood that were processed for previous scales. We adopted a simple implementation based on kdtrees as the underlying data structure of KNNGs, that allows to semydinamically update the graph [18]. Although its theoretically expected complexity is worse than in the case of the state of the art MST algorithms ($O(KN + N \log N)$), in practice the updating notably decreases computation time. It must be noted that the kdtree structure designed by Bentley is not prepared for adding points. Only deletion operations are allowed, thus our implementation estimates entropy from s_{max} to s_{min}.

7 Experimental Results

This section shows several experimental results in order to validate our approach, comparing in terms of time complexity and quality the original Kadir and Brady scale saliency algorithm adapted to multidimensional data and our method based on entropic graphs. In the first case, experiments with hyperspectral images with a total of 31 bands are performed in order to show how the two approaches are affected by the increase of dimensionality of the data. In the second case, a summary of our repeatability experiments using a standard database is shown.

7.1 Computation Time Experiments

We tested our algorithm and the original one by Kadir and Brady with the Bristol Hyperspectral Image Database[1] in order to compare how data dimensionality affects their computation time. This database consists of 29 images, composed by 31 spectrally filtered bands with a resolution of 256×256 and 256 grayscale levels. Thus, each pixel is represented by a vector of 31 greylevel values. Our implementation can directly cope with this input. We also used the scale saliency source code published by Kadir[2], after modifying it in order to be able to process more than 3 dimensions.

The two experiments in Fig. 2 compare the mean computation time of both methods for an increasing number of dimensions (from 1 to 31) for all the images of the database. The grayscale values were quantized in the experiments involving Kadir and Brady implementation, due to the fact that in this case the increasing number of dimensions demands an exponentially increase of memory allocation.

[1] http://psy223.psy.bris.ac.uk/hyper/

[2] http://www.robots.ox.ac.uk/~vgg/research/affine/

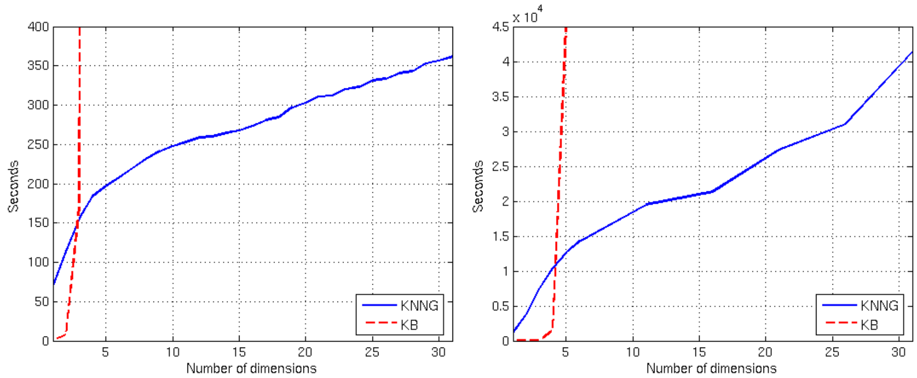

Fig. 2. Comparison of the mean computation time between the original Kadir and Brady scale saliency algorithm (KB) and our approach (KNNG), selecting from 1 to 31 bands of the Bristol database images. Left: $s_{min} = 5$, $s_{max} = 8$, KB grayscale values quantized to 64 histogram bins. Right: $s_{min} = 5$, $s_{max} = 20$, KB grayscale values quantized to 32 histogram bins.

This quantization is not needed when KNNG scale saliency is applied. It is clear from these results that the Kadir and Brady approach is not appropriate for problems where the number of dimensions is higher than 5. On the other hand, the time complexity increase of our approach is linear for higher dimensionalities.

We show results for two different ranges of scales in order to emphasize a limitation of our approach: the width of the range of scales, and especially s_{max}, significantly affects efficiency. As stated before, the KNNG update algorithm used does not allow to add points to the kdtree, only deletion is possible, thus scale space must be computed from s_{max} to s_{min}, building the kdtree for the complete set of points at s_{max} and deleting points through scale space. The inverse route through scale space would be more efficient, but as far as the authors are aware, a general update method with an add operation has not been yet developed.

7.2 Repeatability Experiments

In order to test the quality of the entropy estimation from KNNGs compared to the original Kadir and Brady approach, a complete set of repeatability experiments were performed using the well known dataset used in most feature extractors performance surveys [1]. The evaluation of a feature extraction algorithm is based on the concept of repeatability, a measure of the average number of corresponding regions detected in images under different image conditions. Fig. 3 summarizes some of the results of the repeatability test for several image groups in this dataset, when the 50% of the most salient features are displayed. In all cases, the range of scales spans from $s_{min} = 5$ to $s_{max} = 20$ and the overlap threshold was set to 40% (two regions are considered to match if their ratio of intersection is over 40% after size normalization). Histograms in Kadir

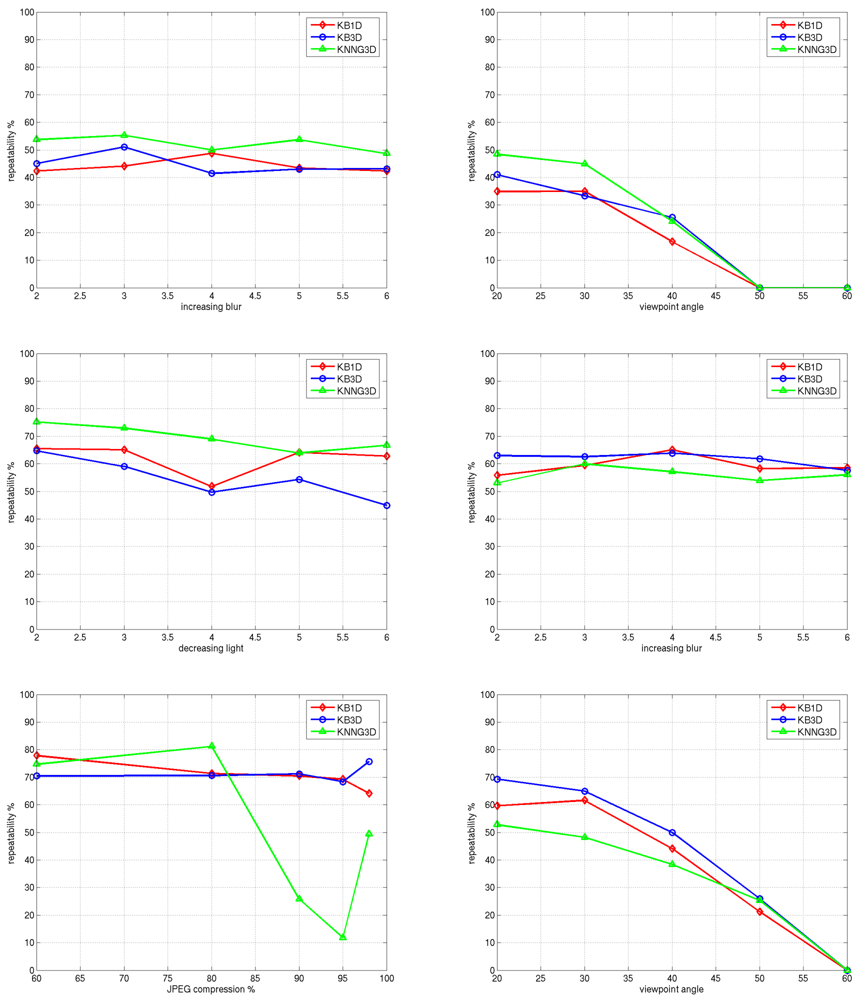

Fig. 3. Repeatability test results for a subset of the image groups in the experiment dataset: *bikes*, *graf*, *leuven*, *trees*, *ubc* and *wall* (group *boat* is a grayscale set of images, and the range of scales used is not appropiate for the *bark* group). These graphs shows the results of three methods: grayscale Kadir and Brady (KB1D), RGB Kadir and Brady (KB3D), and our approach based on KNNGs applied RGB intensity values (KNNG).

and Brady algorithm are quantized to 16 bins (that is the default value of this parameter in the code provided by Kadir and Brady).

The repeatability results of our detector are similar to the ones of the original Kadir and Brady implementation, being slightly better in some cases, with the exception of the ubc sequence. The ubc sequence is aimed to test the repeatability under different JPEG compression rates. As can be seen in Fig. 3, the performance of our detector remarkably decreases for higher compression rates. The main cause is the increase of zero length edges due to the pesence of a higher amount of homogeneous regions, that produces unstability in entropy estimation from the entropic graph [10]. Besides this case, we may conclude that our detector yields similar results in terms of repeatability than the original histogram based implementation when applied to 3 dimensionsal data; thus, althought it can not be proven empirically, we may expect that our algorithm provides equivalent results to those of the original Kadir and Brady algorithm for higher dimensionalities.

8 Conclusions and Future Work

A new high-dimensional scale saliency algorithm based on entropic graphs and inspired by Kadir and Brady algorithm has been presented. Two main steps of the original algorithm, saliency and divergence between scales estimation, are performed by exploiting k nearest neighbors graphs, by means of Rényi α-entropy and Friedman and Rafsky estimator of Henze-Penrose divergence. Our approach has been compared with the original Kadir and Brady algorithm in terms of computation time and quality.

Our future work is aimed to design a better KNNG updating algorithm in order to decrease computation time. Our current approach needs to compute the complete KNNG for each pixel at the maximum scale and then edges must be removed until the minimum scale is reached. The inverse order would be more efficient. We are also considering different practical applications of our algorithm in the fields of hyperspectral and satellite imaging.

References

1. Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T., Gool, L.V.: A comparison of affine region detectors. International Journal on Computer Vision 65, 43–72 (2005)
2. Park, S.J., Shin, J.K., Lee, M.: Biologically inspired saliency map model for bottom-up visual attention. In: Biologically Motivated Computer Vision: Second International Workshop, Tübingen, Germany, pp. 418–426 (2002)
3. van de Weijer, J., Gevers, T.: Boosting saliency in color image features. In: IEEE Conference on Computer Vision and Pattern Recognition, San Diego, USA, pp. 365–372 (2005)
4. Unnikrishnan, R., Hebert, M.: Extracting scale and illuminant invariant regions through color. In: 17th British Machine Vision Conference, Edinburgh, Scotland (2006)
5. Kadir, T., Brady, M.: Saliency, scale and image description. International Journal of Computer Vision 45, 83–105 (2001)
6. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. IEEE Transactions on Pattern Analysis & Machine Intelligence 27, 1615–1630 (2005)
7. Newman, P., Cole, D., Ho, K.: Outdoor slam using visual appearance and laser ranging. In: IEEE International Conference on Robotics and Automation, Florida, USA, pp. 1180–1187 (2006)

8. Fergus, R., Perona, P., Zisserman, A.: Object class recognition by unsupervised scale-invariant learning. In: IEEE Conference on Computer Vision and Pattern Recognition, Wisconsin, USA, pp. 261–264 (2003)
9. Peñalver, A., Escolano, F., Sáez, J.M.: Ebem: An entropy-based em algorithm for gaussian mixture models. In: 18th International Conference on Pattern Recognition, Hong Kong, August 20-24, vol. (2), pp. 451–455 (2006)
10. Neemuchwala, H., Hero, A., Zabuawala, S., Carson, P.: Image registration methods in high-dimensional space. International Journal of Imaging Systems and Technology 16, 130–145 (2006)
11. Hero, A.O., Michel, O.: Applications of spanning entropic graphs. IEEE Signal Processing Magazine 19, 85–95 (2002)
12. Gilles, S.: Robust Description and Matching of Images. PhD thesis, University of Oxford (1998)
13. Kadir, T., Zisserman, A., Brady, M.: An affine invariant salient region detector. In: 8th European Conference on Computer Vision, Prague, Czech Republic, May 11-14, vol. (1), pp. 228–241 (2004)
14. Kadir, T., Boukerroui, D., Brady, M.: An analysis of the scale saliency algorithm. Technical report, University of Oxford (2003)
15. Neemuchwala, H., Hero, A., Carson, P.: Image registration methods in high-dimensional space. International Jounal of Imaging Systems and Technology 16, 130–145 (2006)
16. Henze, N., Penrose, M.: On the multivariate runs test. Annals of statistics 27, 290–298 (1999)
17. Friedman, J.H., Rafsky, L.C.: Multivariate generalizations of the wald-wolfowitz and smirnov two-sample tests. Annals of statistics 7, 697–717 (1979)
18. Bentley, J.L.: K-d trees for semidynamic point sets. In: 6th Annual ACM Symposium on Computational Geometry, California, USA, pp. 187–197 (1990)