Color Invariant Feature Matching For Image Geometric Correction

Aleksandr Setkov
LIMSI-CNRS & Univ. Paris-Sud
BP 133
91400 Orsay, France
aleksandr.setkov@limsi.fr

Michèle Gouiffès
LIMSI-CNRS & Univ. Paris-Sud
BP 133
91400 Orsay, France
michele.gouiffes@u-psud.fr

Christian Jacquemin
LIMSI-CNRS & Univ. Paris-Sud
BP 133
91400 Orsay, France
christian.jacquemin@limsi.fr

ABSTRACT

The success of matching algorithms relies on the definition of features which are both invariant against the geometric distortions to be considered, and distinctive enough to avoid ambiguities. This paper addresses the problem of color feature points matching under photometric and geometric changes. Considering the popular SURF descriptor, it analyzes its state-of-the-art color versions, and proposes a new extension by using local histogram equalization (LHE). While most existing descriptors stem from color conversions and apply to standard lighting variations acquired by the same device, the proposed feature is device-independent and could fit to very generic changes.
The experimental results show that the proposed color descriptors outperform the existing ones under some types of distortions, and are more precise and invariant to different color variations. The paper considers Projector-based Augmented Reality (PAR) as an application field, where one of the evaluation criteria is homography accuracy between real and estimated distorted images. The results show that the proposed method gives the most stable results over all the other techniques and therefore they justify its use for robust color feature matching and its application to geometric correction.

Categories and Subject Descriptors

I.4.7 [**Feature Measurement**]: Feature Representation, Invariants; H.5.1 [**Multimedia Information Systems**]: Artificial, augmented, and virtual realities

General Terms

Theory, Experimentations, Algorithms

Keywords

Color Invariance, Color Feature Matching, Histogram Equalization, SURF, Homography Estimation

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
Copyright 2013 ACM 978-1-4503-2023-8/13/06 ...$15.00.

1. INTRODUCTION

Feature matching algorithms are subjects of many applications in computer vision, robotics, image and stereo processing. Intensity-based methods have appeared in 80s and received a new incentive with appearance of the SIFT algorithm developed by Lowe [10]. Its success caused emergence of many different modifications and alternatives aimed to improve their quality or to reduce their computation time. In these methods, invariance is introduced with respect to geometrical variations such as rotation, translation, scaling, and affine/projective transformations. Furthermore, such descriptors offer good robustness to partial appearances [9, 12, 14]. However working only with grayscale images neglects an important source of information which is represented by the captured color of objects. Some objects can be easily discriminated by the color whilst in grayscale they look similar. Research studies have recently started to combine the two main sources of information : color and geometry and accordingly to introduce color and geometric invariant descriptors for matching. In other words, color invariance can only be built on the basis of image color information. Van de Sande *et al.* [19] provide a good taxonomy of different color spaces which are used for computing descriptors. Among them are HSV, Opponent color space and Weighted opponent color space. Verma *et al.* [20] use in their experiments recently proposed oRGB color space.

This paper also addresses the problem of color feature matching for Projector-based Augmented Reality (PAR) applications. PAR methods aim to achieve a smart projection by compensating radiometric and geometrical distortions (multi-projection is outside the scope of our work). To that purpose, projection devices are enhanced by sensors, for instance a camera to gain information about the environment. A lot of research has been done in this field [3, 1, 21, 15] to efficiently calibrate the system, to compensate different distortions, and to perform 3D registration in order to achieve better quality of final projection onto real-world objects. For example structured light methods rely on matching processes to find correspondence between points in the projected and the captured images [16, 17, 21].

Finlayson *et al.* [5] describe how a well-known histogram equalization technique can be used to provide color invariance. The authors state that the equalization, applied on each color channel independently, has the property to pre-

serve relative color ranking. Therefore, it is invariant to any monotonous color distortion and is independent of the sensor device, which is an important property, required for image indexing. However, this interesting transform is very sensitive to geometric changes, because the device-independance is reached only when the contents of the image is constant. This paper proposes a solution to combine histogram equalization with a local feature point descriptor.

It is important to consider the illumination model under which color invariants have been defined. In this paper two illumination models are addressed : the linear illumination transform used in the taxonomy by van de Sande *et al.* [19], and the non-linear *gamma* function which is typical for most camera sensors.

The remainder of the paper is organized as follows. Section 2 presents the two color changing models and the color invariance models involved in the study. Section 3 describes Feature matching methods used to make correspondences between color invariant features. Section 4 describes the experiments, the evaluation criteria, and the matching results. Finally, section 5 gives a summary of the results and discusses future works.

2. COLOR INVARIANT DESCRIPTORS

In order to define color invariant features, it is interesting to analyze and to model the color variations due to camera settings or changing illumination conditions. The color captured by a camera results from the integration along the visible spectrum wavelengths of several functions which are difficult to compute : the spectrum of lighting sources, the reflectance of the surfaces, and the camera response function for each channel. Unfortunately the physical equation is too difficult to be used directly for color change modeling, unless some simplifying assumptions are made, as in Shafer dichromatic model [18]. Two photometric models are explained in section 2.1. They describe the distortions appearing in images acquired with different camera settings and under different illumination conditions. A few color invariance spaces are mentioned in part 2.2. They are defined so as to have constant values under photometric transformations. Finally, a brief description of the properties of Histogram Equalization is made. As explained further in this paper, the use of this method is appropriate for dealing with the considered distortions.

2.1 Color Change Models

First, let us consider the diagonal model of illumination used by van de Sande *et al.* [19] which defines illuminant color changes and shifts in the resulting image. The equation for this transform is described below :

$$\begin{pmatrix} R^c \\ G^c \\ B^c \end{pmatrix} = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} + \begin{pmatrix} o_1 \\ o_2 \\ o_3 \end{pmatrix} \quad (1)$$

where (R_c, G_c, B_c) are colors after illumination variations, (R, G, B) are colors under the canonical illuminant. The multiplicative parameters a, b, c vary under contrast changes, and o_1, o_2, o_3 model the color shifts.

The second model under consideration is the *gamma* color transform. Most camera sensors have a non-linear response function which relates irradiance of the scene with image brightness. Grossberg and Nayar [8] have collected a database of camera response functions [1] and shown that they all have an exponential form. As shown by many papers on camera-projector systems [15, 4], the projector has very similar non-linear response function. Thus, each response can be modeled by a *gamma* curve expressed as :

$$(R_c, G_c, B_c) = \alpha(R^\gamma, G^\gamma, B^\gamma) \quad (2)$$

where c indicates color distortions that pixels undergo, γ is the exponent. Here (α, γ) are the parameters of the color change model, and are the same for each color channel. α can vary under contrast change.

Starting from these models, it becomes easier to deduce colors invariant features, or to study the invariance property of existing features.

2.2 Color Invariance

Color invariance can be achieved by converting the initial RGB components towards another color space and use this representation for computing descriptors. Let us consider the color transforms described in [19] and histogram equalization from [5]. These invariants are used further in section 3 for feature matching.

Opponent color space. This color space is a linear transform of R,G and B channels. Intensity is represented by O_1 and color information by O_2 and O_3.

$$(O_1, O_2, O_3) = \left(\frac{R-G}{\sqrt{2}}, \frac{R+G-2B}{\sqrt{2}}, \frac{R+G+B}{\sqrt{2}}\right) \quad (3)$$

This color representation is invariant against gray shifts changes, *i.e.* when $o_1 = o_2 = o_3$ in (1) but does not provide invariance against variations in α and γ in (2).

HSV color space. HSV is a well-known perceptual color space which separates chrominance from luminance. Since Hue is a function of a ratio of RGB differences, it is robust against gray shifts changes when $o_1 = o_2 = o_3$, and gray contrast changes when $a = b = c$ in (1), and it is also robust to α variations in (2).

Histogram equalization. Histogram equalization (HE) is a technique used to increase the contrast of grayscale images. When applied to color images, it produces unnatural and distorted colors, so its use for color images has long been argued. For matching purposes, the problem is not related to appearance but to discriminative power and color invariance. As shown in [5], HE preserves rank ordering and is therefore invariant to monotonous color changes.

An HE algorithm transforms image pixels in such a way that the resulting intensity histogram is close to the uniform distribution. It is achieved by computing the cumulative distribution function (*cdf*) and making it linear across the intensity range. The *cdf* should be normalized to $[0, 255]$. Intensity values after histogram equalization appear as follows :

$$I'(i) = cdf(i) \cdot (max(I) - min(I)) + min(I)) \quad (4)$$

where I is the grayscale image, $cdf(i)$ is the cumulative distribution function of i_{th} pixel, operations $max(I), min(I)$ are the maximum and the minimum intensity values in the image.

This method is described here for a grayscale image but it can be easily extended to color by applying it to each channel. Considering the color change models of (1) and (2),

1. This database can be downloaded at `http://www.cs.columbia.edu/CAVE`

both of them are monotonous functions, so HE should handle these distortions very well.

Note that many colorspaces are not considered in our work, either because they provide no invariance when compared to RGB (*e.g.* XYZ), or because they have already been tested for point description, with little success, *e.g.* normalized rgb (L_1 and L_2 of [7]) due to a loss in discriminative power for non-saturated colors as it is the case for C-SURF. This section has considered two models for illumination changes and a few color invariance models that are used in our experiments for descriptor computation.

3. COLOR FEATURE MATCHING

This section first discusses the local feature matching process which comprises three main stages : regions of interest detection, descriptors computation, and feature matching. After that, the Local Histogram Equalization (LHE) method is introduced and its use is justified for Feature Matching. In addition, a discussion is made concerning the validity of LHE for PAR applications.

3.1 Feature Matching

A feature detector extracts the most prominent parts of the image (edges, corners, blobs) to use them as corresponding anchors in a pair of transformed images of the same source. It is shown in [14] that the Hessian matrix-based methods [13] produce better results than the other methods. Detected regions in this case are invariant both to rotation and scale changes. According to the work of Mikolajczyk and Schmid it is better to use Hessian matrix rather than its trace (the Laplacian) for scale selection. As shown in the SURF method [2], Hessian matrix can be efficiently approximated with small degradation in accuracy.

The second stage is descriptor construction for detected feature points. SIFT-like descriptors are some of the best performing and the most widely-used descriptors nowadays. The SIFT algorithm computes a histogram of local oriented gradients around the detected feature point and stores the bins in a 128-dimensional vector (4×4 histograms each with 8 bins). Because intensities are implicitly normalized in SURF, this descriptor is invariant to uniform contrast and intensity changes. The SURF algorithm in many ways approximates the SIFT and other descriptors. It describes a distribution of Haar wavelet responses within the feature point neighborhood. Integral images are used for speed and only 64 dimensions are used to reduce matching time.

At the second stage the problem of color invariance representation becomes important, and it is necessary to make descriptors invariant not only to geometric but to illumination changes. For these purposes the initial color space is pretransformed to make descriptor representation independent of illumination changes. As for color transformations, one of the models described in section 2 can be applied. A local color feature descriptor can be constructed by concatenation of intensity-based descriptors computed independently for each channel. As a result, a large descriptor is obtained which is further used in matching. According to the work of Van de Sande *et al.* [19], RGB- and Opponent-SIFT produce very good matching quality and they are among the methods with the best overall performance in the evaluation.

The last stage is descriptor matching. In this process each feature descriptor in one image is compared with each descriptor in another image. There exist several ways to define a match. One way is a threshold-based matching where features are matched if the distance is below a threshold. Each feature therefore can have several matches. Another way is nearest neighbor-based matching where two descriptors are considered as a matched pair if the distance between them is below one threshold and the distance to the second closest descriptor is greater than another threshold. In this case each descriptor can have at most only one match. In our experiments we will use the second strategy, because in most applications each point can only have one correct match.

3.2 Local Histogram Equalization for Feature Matching

As explained in section 2.2, HE can provide invariance against any monotonous illumination change, therefore it is more generic than most existing color invariants. However, the transform is generally made globally in the image, because the invariance holds only when the contents of the two images to be matched are perfectly similar. Otherwise the rank preservation cannot be ensured. This is unfortunately a very important constraint. On the other hand, color variations are not always similar in all the areas of the scene, and the global HE (noted GHE in the rest of the paper) would provide poor results in this situation. The question is then : how to achieve Local HE (noted LHE) while being sure of the stability of the image contents ?

Most local feature descriptors have the advantages to require only local color information around each feature point, and to be very stable under geometric distortions (local image contents is preserved). Then, once all points are precisely detected after the first stage, it is easy to extract rectangular regions centered around these points according to their orientation and scale. Then, LHE can simply be performed on this region in a reliable way. For sake of clarity, Fig. 1 illustrates the process. The detection of the points can be made in the grayscale image, and the LHE performed on each color channel before matching. More than being a color invariant for the matching of several images acquired by the same sensor, LHE is also well suited for camera-projector systems. This assertion is discussed in the following section.

3.3 Discussion on local histogram equalization for PAR applications

Let P be a physical point in the scene (real world) lighted by a projector which produces an energy depending on its input image at a pixel p, $A_k(p)$, where k is the color channel R,G, or B. In the general case, the viewed color component $B_k(q)$ is produced by the integration along the wavelengths of : 1) the illuminant spectrum $\mathcal{E}(\lambda, q)$, which is the emission spectrum produced by the projector, 2) the surface reflectance $R(\lambda, q)$ which defines the proportion of incident light which is reflected by the surface for each λ ; 3) the spectral sensitivity of the sensor $S_k(\lambda)$ for the k^{th} channel. Assuming that there is no ambient lighting in the scene, then :

$$B_k(q) = \int_\lambda \mathcal{E}(\lambda, q)\mathcal{R}(\lambda, q)\mathcal{S}_k(\lambda)d\lambda \qquad (5)$$

Note that p is the initial location of the point in A_k, and q is its location in the distorted image B_k. The spectrum $\mathcal{E}(\lambda, q)$ produced by a projector depends on the technology,

which is generally tri-LCD [2] or DLP [3]. In the case of tri-LCD, the video signal is decomposed into three components. Each color beam illuminates a monochrome LCD panel, then the three channels are merged *via* a prism. The lamps are based on mercury vapors, the emission spectrum is discrete and has generally three monochromatic beams, sometimes more. Therefore $\mathcal{E}(\lambda, p) = 0$ except on some wavelengths. Thus, in the case of one wavelength per channel k, (5) can be simplified as :

$$B_k(q) = \mathcal{E}(\lambda_k, q)\mathcal{R}(\lambda_k, q)\mathcal{S}_k(\lambda_k) \quad (6)$$

In that case, the energy at the wavelength λ_k is a monotonic function F_k^v of the original intensity in that bandwidth, generally a gamma function as said previously :

$$\mathcal{E}(\lambda_k, q) = F_k^v(A_k(p)) \quad (7)$$

Finally, under these assumptions, equation (5) becomes :

$$B_k(q) = F_k^v(A_k(p))\mathcal{R}(\lambda_k, q)\mathcal{S}_k \quad (8)$$

where S_k is finally a constant related to the sensor gain, but can vary slowly during time due to the aging of the system. Let us now discuss on the validity of the rank conservation, with respect to the type of projection surface.

White Lambertian surfaces. Consider two pixels q and q' localized on the same surface. When the surface is white and Lambertian then $\mathcal{R}(\lambda_k, q) = 1$ and :

$$B_k(q) = F_k(A_k(p))\mathcal{S}_k$$

Therefore when $A_k(p) < A_k(p')$ then $F_k(A_k(p)) < F_k(A_k(p'))$ and $B_k(q) < B_k(q')$. The ranks are preserved and the HE proposed by Finlayson is a color invariant in that context.

Uniform surfaces. When $\mathcal{R}(\lambda_k, q) = \mathcal{R}(\lambda_k, q')$, $\forall q, q'$, which implies that the surface is Lambertian, then if $A_k(p) < A_k(p')$ then $F_k(A_k(p)) < f_k(A_k(p'))$ and $B_k(q) < B_k(q')$. On a same Lambertian projection surface, the rank of colors is preserved between the ideal image A_k and the distorted image B_k .

Non uniform surfaces. Consider two neighbor pixels p and p' localized on two areas with different reflectance properties, then the rank is preserved only in some situations. In other words, $B_k(q) < B_k(q')$ when : 1) $A_k(p)$=$A_k(p')$ and $\mathcal{R}(\lambda_k, q) < \mathcal{R}(\lambda_k, q')$; 2) $A_k(p) < A_k(p')$ and $\mathcal{R}(\lambda_k, q) < \mathcal{R}(\lambda_k, q')$; 3) $A_k(p) < A_k(p')$ and $\mathcal{R}(\lambda_k, q) = \mathcal{R}(\lambda_k, q')$. It yields that, for non uniform surfaces, the rank is preserved only locally, either when the region is uniform in terms of colors or when the colors of both image and surface vary in the same way, *i.e.* when both of them increase or decrease. By assuming that the surface vary smoothly in terms of reflectance $\mathcal{R}$, we can assume that its projection on the image is locally constant in a small neighborhood W, *i.e.* $\mathcal{R}(\lambda_k, q) = \mathcal{R}(\lambda_k, q')\forall q, q' \in W$. Then the color rank is preserved locally in W.

Thus, the GHE (*i.e* made once in the whole image) provides color invariance only when the projection surfaces are of uniform reflectance everywhere in the scene under consideration. The LHE provides color invariance also when these surfaces have constant reflectance locally. Note also that the invariance is not guaranteed when the feature point is located on the edges between surfaces of very different colors,

2. This technology is transmissive and is based on tiny and transparent LCD screens (0,55" to 0,9")

3. DLP technology is reflexive and is based on thousand tiny mobile mirrors

Figure 1: Example of LHE for a feature point

for example between two tapestries. When *a priori* knowledge is available on both the image to be projected and on surface properties, it is possible to assess whether the size of the neighborhood chosen for LHE is appropriate for color rank preservation.

The experiments made in the following section analyze the behavior of the proposed LHE SURF matching, when compared to existing methods.

4. EXPERIMENTAL RESULTS

This section describes the experiments made to evaluate different color invariance descriptors under various photometric and geometric distortions. First, the experimental setup is described. Next we review the evaluation criteria used for descriptor assessment. Here we consider both general feature matching performance and homography estimation as the principal tasks of geometric compensation algorithms for PAR. Finally, the experimental results for all the considered methods are shown.

4.1 Experimental Setup

Figure 2 represents the processing pipeline of the experiments being conducted in the framework of our research work. It shows the main steps that we perform to evaluate color invariant descriptors. We start with an RGB test image of 256 by 256 pixels resolution. For this image three classes of distortions are introduced : photometric, geometric, and both color and geometric, which are noted *Photo.*, *Geo.* and *Photo. + Geo* respectively (some examples can be seen in figure 3). For each class, $N = 100$ transformations are synthesized and applied to the initial image. As a result, the generated test database consists of $2N$ images with photometric distortions, N images with geometric distortions and $2N$ images with both photometric and geometric distortions. As for photometric distortions, the two illumination change models of 2.1 are applied. An example of such a distorted image is presented as *step 1* in the figure 2. We randomly choose parameters for the linear and the *gamma* color transforms. As for geometric transformations we generate random homography matrices and use them to generate various warpings of the test image. We changed the coordinates of the four image vertices and used a standard OpenCV function to get a homography matrix. The transformation radius was chosen randomly to be approximately equal to one third of the image size at maximum.

The next step is feature detection (*step 2*). The feature points are the same in each experiment in order to provide a fair comparison. To that purpose, they are extracted in the luminance images. As the detected points are the same for each considered color space, we can focus on comparing

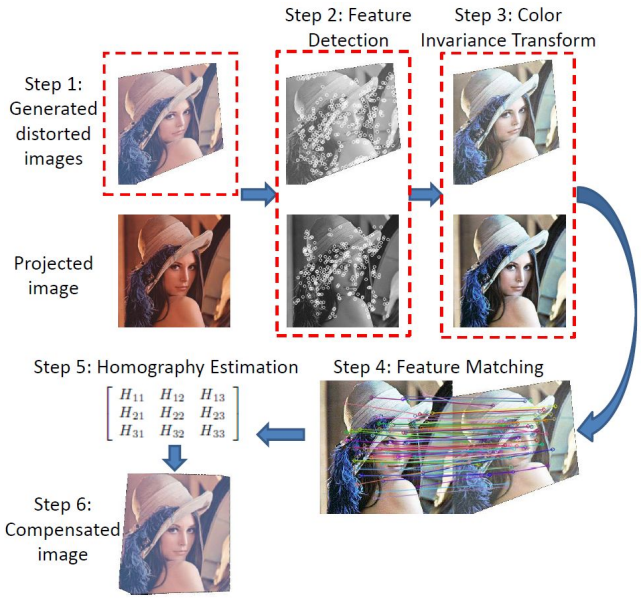

Figure 2: Processing pipeline of Color Homography Estimation Experiments

descriptions, which is the main objective of the paper. After that, we perform color invariance transformation on the test and distorted images (*step 3*) and pass to descriptors computation and feature matching (*step 4*).

The SURF method was used in our experiments because of its successful use in different computer vision applications. SURF is similar to SIFT in terms of results and nature of operations. It only uses an approximate version of the operations used in SIFT and therefore is several times faster. The processing time comparison, done by L. Juan and O. Gwun in [11], shows that SURF feature matching is almost four orders faster than SIFT. Note however that the study could also hold for SIFT. We used the SURF implementation of E. Oyallon and J. Rabin available at [4]. The parameters were chosen to give the best matching performance on a set of 100 test images. They were fixed and kept unchanged throughout the experiments. Except for the implementation of SURF, we used the OpenCV library for image processing and homography estimation. Note that although OpenCV is a well-optimized tool, the purpose of this paper is to analyze quality of the algorithms rather than their runtime performances.

The following point descriptors are used :

1. SURF is the classical SURF computed on the luminance I ;
2. LHE is the classical SURF with local histogram equalization ;
3. GHE is the classical SURF with global histogram equalization, made on the whole image ;
4. RGB is the color feature SURF with a descriptor of size 68×3 ;
5. RGB + LHE is the RGB SURF for which LHE has been performed on each channel ;
6. HSV is the same implementation of RGB SURF with a conversion to HSV ;
7. $O_1O_2O_3$ or Opponent SURF, is similar with conversion to $O_1O_2O_3$ colorspace.

Note that C-SURF, where O_1 and O_2 are divided by O_3, is not used in the evaluation. The first reason is that our experiments have shown very unstable results, probably because the descriptor distinctiveness is reduced. In addition, invariance against intensity is already implicitly performed in SURF, as mentioned in 3.1.

The two final stages (*steps 5,6*) correspond to homography estimation and geometric compensation. They will be described in detail in the next section.

4.2 Evaluation Criteria

We examine the impact of color descriptions on three main properties : invariance, distinctiveness, and precision.

Color invariance. The first property shows how matching quality deteriorates when introducing geometric and/or color distortions. Concerning geometric distortions, invariance is provided directly by the SURF procedure, which has been designed to offer invariance to rotation and scale. Therefore, the geometric invariance should be the same for each method. We evaluate the quality of color descriptors by comparing their performances under various distortions. More precisely, for each test image we compute the normalized number of correct matches and then we average it over the number of test images. This criterion CMR, for Correct Matches Ratio, is used to evaluate descriptor invariance :

$$CMR = \frac{\sum_{i=0}^{N-1} \frac{\#\text{ correct matches }(i)}{\#\text{ total matches}(i)}}{N} \tag{9}$$

N is the number of test images, $\#$ correct matches (i) and $\#$ total matches (i) are the number of correct matches and the total number of matches in image i respectively.

Distinctiveness. This criterion defines a possibility to discriminate objects that are represented by feature descriptors. It ensures that, for each feature point, there is only one matched point that produces a distance significantly lower than the second nearest point does. We examine descriptors on this property by computing the total number of correct matches divided by the total number of matches over all the test images. This metrics shows how the number of correct matches CM varies after applying the distortions defined in section 4.1.

$$\#CM = \frac{\sum_{i=0}^{N} \#\text{correct matches}(i)}{\sum_{i=0}^{N} \#\text{total matches}(i)} \tag{10}$$

Precision. In order to estimate the precision of the matching, let us consider a registration problem. In PAR applications, this task will be useful to compensate geometric distortions of the projected video. Once feature matches are computed for the projected and the captured distorted image, the projective transform between the images can be estimated. It can be done by several different methods but we chose the RANSAC algorithm [6] which is implemented in the OpenCV library. It is one of the most used and best performing algorithms for this purpose.

To compare homography matrices, we first apply the inverse randomly generated homography transform $\boldsymbol{H_r}$ to the the distorted image. We do the same operation with the inverse estimated homography matrix $\boldsymbol{H_e}$. Homogeneous coordinates of these warped images are transformed to Cartesian

4. `http://www.ipol.im/pub/pre/69/`

Figure 3: Examples of distorted images. The first row represents geometric transforms, the second and the third rows - gamma + geometry and linear color + geometry transformations respectively.

X_e, Y_e, X_r, Y_r :

$$(X_e\ Y_e) = (x_e/z_e\ y_e/z_e) \quad (11)$$
$$(X_r\ Y_r) = (x_r/z_r\ y_r/z_r)$$

where $x_e, y_e, z_e, x_e, y_e, z_e$ are the coordinates of the distorted image warped by $\boldsymbol{H_e}$ and $\boldsymbol{H_r}$.

Finally, the metrics to estimate how close the compensation images are in the geometric sense can be computed as follows :

$$d(H_e, H_r) = \sum_{\substack{X_e=0,Y_e=0 \\ X_r=0,Y_r=0}}^{W,H} \sqrt{(X_e - X_r)^2 + (Y_e - Y_r)^2} \quad (12)$$

where W, H are the width and height of all the images; $\boldsymbol{H_e}$ and $\boldsymbol{H_r}$ are 3×3 estimated and real inverse homography matrices. We assume, that the lower the distance, the higher the precision of the global feature points matching. This criterion is analyzed in section 4.4 on synthetic and real images.

4.3 Experimental Results

First, we consider the results when only photometric distortions are applied to the test image. In this case there are no geometric transformations, and it is easy to check true matches by comparing feature point coordinates. The first columns of table 1 and table 2 show matching performance of the descriptors under linear color changes. According to these results the use of RGB SURF with LHE gives the best performance among all the descriptors. Under *gamma* color change, Opponent SURF slightly outperforms this method (the second columns of the same tables).

The third columns of tables 1 and 2 correspond to the matching results under geometric distortions. As the real homography matrix is known, we can apply this transformation to the coordinates of the feature points to estimate ground-truth matches and consequently to compute exactly the same metrics as for photometric distortions. According to the results HSV-SURF works very well and produces the best performance. That means that the use of this descriptor is reasonable if only geometric distortions occur.

The last two columns provide us with the matching results when both color and geometric distortions are applied simultaneously. It is important to mention that both HSV- and Opponent-SURF produce significantly poor results under linear color distortions coupled with geometric transforms. As mentioned in 2.2 HSV colorspace is robust only to grayscale photometric changes. Although Opponent-SURF yields slightly better results in the case of gamma color and geometric transformations, RGB with LHE SURF gives more stable results without considerable losses in quality for each distortion scenario. It is worth to mention that, since photometric changes are uniform in the whole image, it makes the use of GHE favorable. It is proved by the results for which GHE SURF is more accurate than Intensity-based SURF (see tables 1 and 2).

4.4 Color Feature Matching Results for PAR Applications

This subsection shows results of color feature matching for homography estimation. Two cases are considered here. The first one corresponds to the tests on synthetic images from the previous subsection. We focus on the comparison of homography matrices estimated on the basis of feature matching with ground-truth homographies. The second case corresponds to real images. As we do not have ground truth homographies to compare with, we show the results for user-based quality evaluation.

Table 3 shows the performances of the descriptors according to the estimated homographies. As explained in section 4.2 we use Euclidean distance between compensated coordinates as a measure of error with respect to exact coordinates. The very high error values in the table for HSV- and Opponent-SURF can be explained by the fact that, on some test images, these methods fail in computing enough correct matches to correctly estimate the homography transformation. For that reason they produce large error values which are accumulated in the final mean error value. Although RGB with LHE SURF is not the best performing method in case of geometric only transformations, it produces robust results for all the three experimental distortion combinations.

In figure 4 the real compensated images are presented. The captured image was taken under geometric and lighting distortions. Also it can be seen that the background surface is not single-color and this fact makes the matching process even harder. In this example we consider only geometric compensation made by means of the estimated homography. In the figure the results of all considered methods are shown. We can see that except for HSV-SURF based compensation all the images look similar. The Grayscale-SURF with LHE, the RGB-SURF, and the GHE compensated images seem to be closer to the correct result. This proves the fact that using LHE can enhance quality of compensated images.

5. CONCLUSIONS

Based on the review on existing illumination change models and color feature descriptors, the paper has proposed a new SURF descriptor based on local histogram equalization. In addition to its intrinsic geometric invariance, it has the property to be invariant to monotonous illumina-

Figure 4: Real homography transform estimation experiments. (a) Projected image ; (b) Captured image ; The following images compensated by : (c) SURF ; (d) SURF with LHE ; (e) RGB SURF ; (f) RGB SURF with LHE ; (g) SURF with GHE ; (h) HSV SURF ; (i) Opponent SURF.

Table 1: Correct Matches Ratio in % (see section 4.2) for different SURF descriptors under geometric and photometric variations.

Feature	Photo		Geo	Photo+Geo	
	Linear	Gamma		Linear	Gamma
RGB	96.18	97.63	84.66	82.97	84.30
RGB + LHE	**97.05**	98.03	86.03	**84.99**	86.52
I	95.72	97.18	84.13	80.25	81.67
I + LHE	96.03	97.44	81.75	81.33	82.80
HSV	90.15	79.16	**89.09**	69.93	67.02
$O_1O_2O_3$	93.2	**98.22**	87.83	74.08	**87.13**
GHE	96.93	97.85	84.71	83.38	84.84

Table 2: Number of Correct Matches in % (see section 4.2) for different SURF descriptors under geometric and photometric variations.

Feature	Photo		Geo	Photo+Geo	
	Linear	Gamma		Linear	Gamma
RGB	96.67	97.88	86.68	84.95	85.93
RGB + LHE	**97.50**	98.16	88.28	**87.35**	88.17
I	96.28	97.50	84.13	82.57	83.41
I + LHE	96.56	97.66	86.85	83.70	84.83
HSV	94.65	95.54	**90.50**	75.94	81.76
$O_1O_2O_3$	94.75	**98.30**	89.26	78.90	**88.69**
GHE	97.34	98.0	86.85	85.51	86.39

tion changes, and to be device-independent. Owing to our study, LHE can be used for PAR applications, where the distortions to be compensated are complex, because they are due to the projection on a non-ideal surface and the acquisition by the camera. Experiments have been performed on synthetic tests and real images, and three criteria have been considered. Contrary to most color invariants, the proposed method produces stable and robust results for all tested photometric and geometric distortions. The best performance is achieved under *gamma* photometric changes where it outperforms all the other methods. LHE has been successfully used for homography estimation, which is the standard geometrical transformation used in camera/projector systems.

Computation time optimization issue comes up when using LHE for real-time color feature matching. In our experiments we used a non-optimized implementation which takes more processing time than RGB SURF because LHE is performed in each region independently. To be used in real-time applications the algorithm requires optimization that can be obtained through parallel processing of feature regions.

The paper considered two photometric distortion models. Although they are easy-to-model and widely used in literature, they are not necessarily good approximations of PAR physical conditions. In this way some investigation is required to construct a more precise model that better fits these conditions and that can be easily addressed by color invariant methods. In future works, we will also go further in

Table 3: Average Euclidean Distance (see section 4.4 for the different SURF descriptors under geometric and photometric variations.

Feature	Geo.	Linear + Geo.	Gamma + Geo.
RGB	1.69	2.38	1.83
RGB + LHE	1.55	**1.90**	1.70
I	1.54	2.06	**1.61**
I + LHE	1.69	2.0	1.79
HSV	**0.95**	117.20	108.45
$O_1O_2O_3$	1.57	35.18	1.67
GHE	1.65	2.07	1.63

the experiments to assess the device-independence property of the descriptor.

6. REFERENCES

[1] M. Bajura and U. Neumann. Dynamic registration correction in video-based augmented reality systems. *IEEE Comp. Graph. and Appli.*, 15 :52–60, 1995.

[2] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. Speeded-up robust features (surf). *Comput. Vis. Image Underst.*, 110(3) :346–359, June 2008.

[3] O. Bimber and R. Raskar. *Spatial Augmented Reality : Merging Real and Virtual Worlds*. A. K. Peters, Ltd., Natick, MA, USA, 2005.

[4] J. Dehos, E. Zeghers, C. Renaud, F. Rousselle, and L. Sarry. Radiometric compensation for a low-cost immersive projection system. In *Proc. of ACM symposium on Virtual reality softw. and techno.*, VRST '08, pages 130–133, New York, NY, USA, 2008. ACM.

[5] G. D. Finlayson, S. D. Hordley, G. Schaefer, and G. Y. Tian. Illuminant and device invariant colour using histogram equalisation. *Pattern Recognition*, 38 :179–190, 2005.

[6] M. A. Fischler and R. C. Bolles. Random sample consensus : a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6) :381–395, 1981.

[7] T. Gevers and A. W. M. Smeulders. Colour based object recognition. *Pattern Recognition*, 32(3) :453–464, March 1999.

[8] M. Grossberg and S. Nayar. What is the Space of Camera Response Functions ? In *IEEE CVPR*, volume II, pages 602–609, Jun 2003.

[9] D. G. Lowe. Object recognition from local scale-invariant features. In *ICCV*, 1999.

[10] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. Jour. Comput. Vision*, 60(2) :91–110, 2004.

[11] O. G. Luo Juan. A comparison of sift, pca-sift and surf. *Inter. Jour. of Im. Proc. (IJIP)*, 3 :143–152, 2009.

[12] K. Mikolajczyk and C. Schmid. An affine invariant interest point detector. In *7th ECCV*, pages 128–142. Springer, 2002. Copenhagen.

[13] K. Mikolajczyk and C. Schmid. Scale and affine invariant interest point detectors. *Int. Jour. of Comp. Vision*, 60(1) :63–86, 2004.

[14] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *IEEE Trans. on PAMI*, 27(10) :1615–1630, 2005.

[15] S. Nayar, H. Peri, M. Grossberg, and P. Belhumeur. A Projection System with Radiometric Compensation for Screen Imperfections. In *ICCV Workshop PROCAMS*, Oct 2003.

[16] H. Park, M.-H. Lee, B.-K. Seo, and J.-I. Park. Undistorted projection onto dynamic surface. In *Proc. PSIVT'06*, 2006.

[17] H. Park, M.-H. Lee, B.-K. Seo, J.-I. Park, M.-S. Jeong, T.-S. Park, Y. Lee, and S.-R. Kim. Simultaneous geometric and radiometric adaptation to dynamic surfaces with a mobile projector-camera system. *IEEE Trans. Cir. and Sys. for Video Technol.*, 18(1) :110–115, Jan. 2008.

[18] S. A. Shafer. Color. chapter Using color to separate reflection components, pages 43–51. Jones and Bartlett Publishers, Inc., USA, 1992.

[19] K. E. A. van de Sande, T. Gevers, and C. G. M. Snoek. Evaluating color descriptors for object and scene recognition. *IEEE Trans. on PAMI*, 32(9) :1582–1596, 2010.

[20] A. Verma, S. Banerji, and C. Liu. A new color SIFT descriptor and methods for image category classification. In *Int. Cong. on Comp. Appli. and Comput. Sci.*, pages 819–822, 2010.

[21] S. Zollmann, T. Langlotz, and O. Bimber. Passive-active geometric calibration for view-dependent projections onto arbitrary surfaces.