Neurocomputing 265 (2017) 28–41

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Maximal similarity based region classification method through local image region descriptors and Bhattacharyya coefficient-based distance: Application to horizon line detection using wide-angle camera

Y. El merabet[a,*], Y. Ruichek[b], S. Ghaffarian[d], Z. Samir[a], T. Boujiha[a], R. Messoussi[a], R. Touahni[a], A. Sbihi[c]

[a] Laboratoire LASTID, Département de Physique, Faculté des Sciences, Université Ibn Tofail, B.P 133, Kenitra 14000, Morocco
[b] Le2i UMR 6306, CNRS, Univ. Bourgogne Franche-comté, UTBM, Belfort F-90010, France
[c] Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede 7500 AE, The Netherlands
[d] Laboratoire LABTIC, ENSA, Université Abdelmalek Essadi, Route Ziaten, km 10, Tanger BP 1818, Morocco

ARTICLE INFO

Article history:
Received 15 February 2016
Revised 4 March 2017
Accepted 20 March 2017
Available online 3 June 2017

Keywords:
GNSS
Region classification
Image segmentation
Color invariance
Color texture feature
Hybrid descriptor
Maximal similarity

ABSTRACT

In recent years, many approaches have been proposed to compensate the lack of performance of GNSS (Global Navigation Satellites Systems) occurring when operating in constrained environments. One of these approaches consists in characterizing the environment of reception of GNSS signals using a wide-angle (fisheye) camera oriented to the sky. The content of acquired images is classified into two regions (sky and not-sky) in order to determine LOS (Line-Of-Sight) satellites and NLOS (Nonline-Of-Sight) satellites. This paper is aimed at proposing an image-content classification method to make this approach more effective. The proposed method is composed of four major steps. The first one consists of simplifying the acquired image with an appropriate couple of colorimetric invariant and exponential transform. In the second step, the simplified image is segmented using Statistical Region Merging method. The third step consists of characterizing the segmented regions with a number of local image region descriptors providing more statistically meaningful and discriminatory features. In order to classify the characterized regions into sky and non sky regions, we propose the supervised $\mathcal{MSRC}$ (Maximal Similarity Based Region Classification) method by using Bhattacharyya coefficient-based distance. Comparative and extensive experiments have been conducted to investigate the effectiveness of the proposed $\mathcal{MSRC}$ method according to the proposed groups of local image region descriptors. Furthermore, we clearly validate the feasibility of $\mathcal{MSRC}$ method by comparing its results with those presented in the state of the art.

1. Introduction

Nowadays, GNSS systems (Global Navigation Satellites Systems), such as GPS, European GALILEO, COMPASS and GLONASS contribute widely to localization and navigation systems in Intelligent Transport Systems (ITS). They have penetrated the transport market through applications such as monitoring of containers, fleet management, etc. Although most of them give satisfying accuracy in terms of positioning, they cannot avoid propagation problems, caused by multi-path phenomena (cf. Fig. 1) of GNSS signals (occurring mainly in constraint environments such as urban zones). Indeed, in dense environments, signals can be blocked (no signal received), shadowed (signal received after reflections without any direct ray) and direct. These constraints make tough the evaluation of estimated position reliability. This drawback has no negative impact for applications like containers monitoring, flot management, etc., not requiring high availability, integrity and accuracy of the positioning system. On the contrary, for specific applications dealing with liability issues (toll, insurance, etc.) as well as safety-related applications (automatic guidance or control), requiring more stringent performances, this is a real challenge. There are many techniques for localization performance enhancement in the literature. Multi-sensor-based approaches allowing to compensate the lack of performance of GNSS by adding other sensors (odometer, Inertial Measurement Unit, etc.) that increase the system complexity are commonly used [1–3]. Another workaround consists in using complementarity between computer vision and localization systems to characterize the environment of reception of satellites. Koch and Teller [4] propose a localization method that uses

* Corresponding author. Tel: +212638875312.
E-mail addresses: y.el-merabet@univ-ibntofail.ac.ma (Y. El merabet), yassine.ruichek@utbm.fr (Y. Ruichek).

http://dx.doi.org/10.1016/j.neucom.2017.03.084
0925-2312/

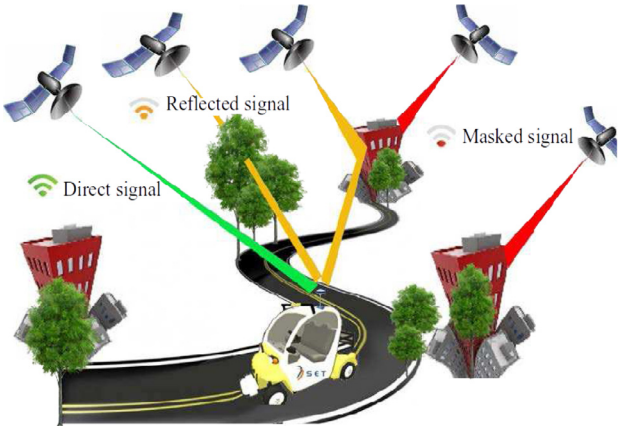

Fig. 1. Illustration of the multipath phenomenon in urban areas.

a known 3D model of the treated environment and a wide angle camera. This method seems inadequate face to the real-time constraint required for the application due to its low frame rate 0.1 image/s. Meguro et al. [5] propose a method that aims to improve the positioning accuracy. The authors use an infrared camera to address the elimination of NLOS satellites, considered as a factor degrading the performance of the positioning system. The detection of the horizon line is mainly based on an azimuthal projection. However, the result does not reflect the real horizon line and the approach can potentially lead to errors when the camera is not vertically oriented. Ramalingam et al. [6] propose a method aiming to detect the horizon line using an omni-directional camera in the visible range, as well as a 3D model of the environment. Another use of 3D models is presented in [7], which compares satellite real visibility and theoretical visibility obtained with the model to refine lateral accuracy of the position, and in [8] which proposes a pseudo-range bias estimation with EKF. Unfortunately, the major drawback of these 3D model-based approaches is that the obtention of a precise 3D model of the environment is not guaranteed.

Recently, like Meguro [5], the authors in [9] propose to enhance localization accuracy by analyzing the structure of the environment traveled by a vehicle using a single camera delivering visible range to overcome the problems previously identified (lack of precision of 3D models, time computation, etc.). The aim is not to develop a new sensor but to study how could be improved the performance of large existing public GNSS receivers from the analysis of images provided by a wide-angle camera (fisheye camera with a large field of view of 180°) mounted near to a GPS receiver on the roof of the mobile and oriented to the sky. The fisheye images are acquired with the GPS acquisition frequency when the vehicle is traveling. Once an image is acquired, two major steps are sequentially applied. The first one, based on image processing, consists of sky region detection. For that, the acquired image is simplified using a geodesic reconstruction with an optimal contrast parameter. Then, a clustering step is performed in order to classify the regions into two classes (sky and not-sky). A set of unsupervised (Km local, Fuzzy Cmeans, Fisher and Statistical Region Merging) and supervised (Bayes, K-Nearest Neighbor and Support Vector Machine) clustering algorithms have been tested and compared. The second step consists of repositioning satellites in classified images to identify GNSS signals with direct path (resp. blocked/reflected signals) i.e. located in sky region of the image (resp. located in not-sky region). More details of this repositioning step can be found in [9]. It might be worth to mention that the reliability of the proposed system depends greatly on the classification results. The work presented in this paper, is related to the first step and tries to make the proposed framework [9] more effective in terms of image classification results.

2. Main contributions

In this paper, we propose to improve the framework introduced in [9] in terms of classification of the content of fisheye images into two class (sky, not sky). Fig. 2 illustrates the flowchart of the image processing-based method for localization. The method is composed of two blocks: image processing and localization. Our contribution is concerned with the image processing part. The method we propose is composed of several steps: (1) image simplification, (2) image segmentation, (3) region features extraction and (4) region classification. These steps will be detailed in next sections. The main contributions of this work are:

- An automatic method for choosing an appropriate couple of colorimetric invariant/image enhancement for image simplification purposes is adopted. The effect of this useful preprocessing step is twofold: (1) it gives more precise segmentation that is faithful to the desired real objects and (2) it allows sufficient attenuation of over-segmentation problem.
- We extended color invariance to the RGB color space and proposed 21 local color invariance histograms using image quantization in order to increase the photometric invariance properties and enhance the discriminative performance.
- We also extended a set of existing texture descriptors to RGB color space and proposed 10 color local texture histograms.
- Based on these local histograms, we built a number of 31 new local image region descriptors, denoted as local hybrid histograms by concatenating color RGB histogram on the one hand with color invariance histograms leading to 21 hybrid color histograms, and on the other hand with color local texture histograms leading to 10 hybrid color texture histograms.
- In order to classify the regions of the segmented images, we proposed a $\mathcal{MSRC}$ (maximal similarity based region classification) method by using Bhattacharyya coefficient-based distance.
- Comparative and extensive experiments have been conducted to investigate the effectiveness of the $\mathcal{MSRC}$ method according to the proposed groups of local color invariance, local color texture and local hybrid image region descriptors. Furthermore, we clearly validate the feasibility of the $\mathcal{MSRC}$ method by comparing its results with those of recent state-of-the-art methods [9,10].

The organization of the paper is as follows: Section 3, details the image simplification step of the proposed procedure. Section 4 presents briefly the Statistical Region Merging algorithm used to obtain the preliminary fisheye image segmentation. In Section 5, we introduce the proposed local color invariance, color local texture and color hybrid histograms. Section 6, presents the proposed $\mathcal{MSRC}$ algorithm. Experimental results and discussions are given in Section 7. Conclusions are derived in Section 8.

3. Image simplification

Image simplification, which reduces content information of an image to suppress undesired details such as noise, is a very important basic pre-processing step of a lot of image-based applications. This step consists of choosing an appropriate couple of colorimetric invariant/image enhancement for image simplification purposes allowing increasing the robustness of sky extraction.

3.1. Colorimetric invariants (CI)

In this paper, we investigate the effect of CI on the outcome of SRM-based fisheye image segmentation. In other words, the objective is to show how using color invariance could limit the artifacts

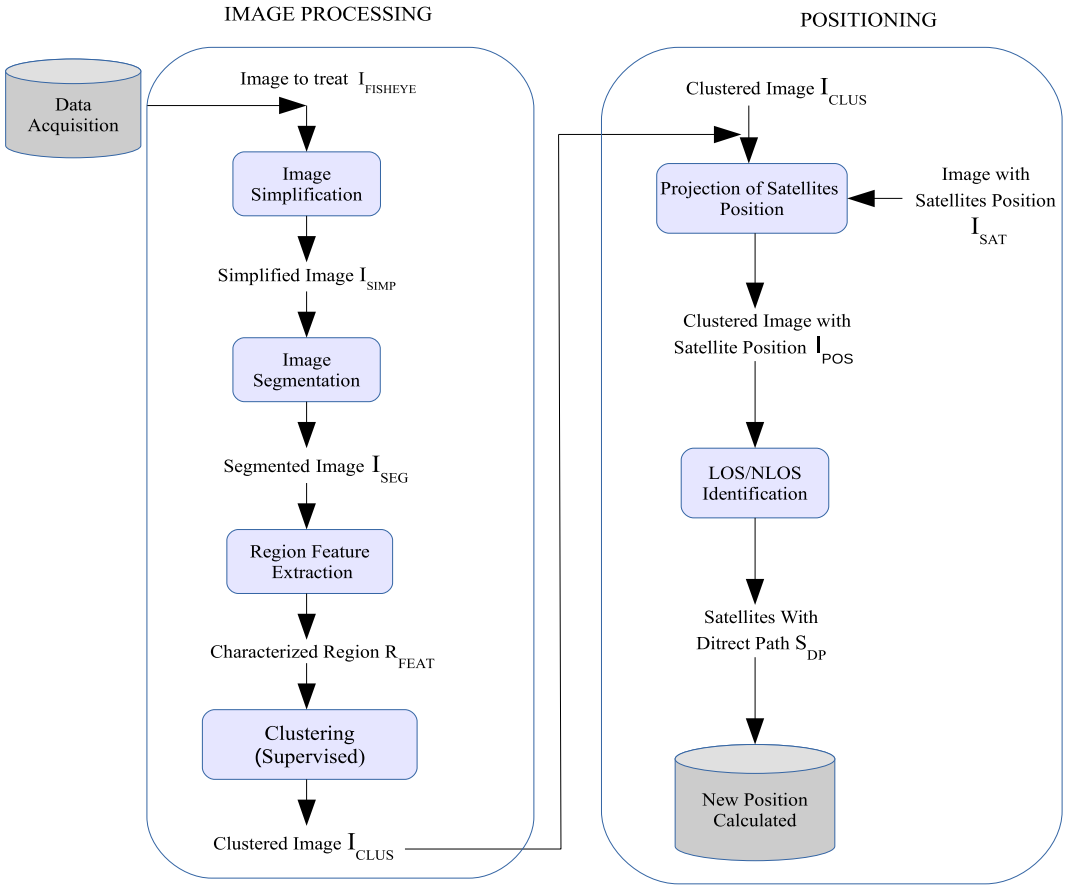

Fig. 2. Flowchart of the image processing-based method for localization.

present in the acquired images (noise and unimportant fine-scale details) and in fine enhance the fisheye image segmentation quality.

In the context of our application, the color of the images acquired in mobility with a mobile platform (instrumented vehicle) is heavily dependent on several factors such as the surface reflectance, illuminant color, lighting geometry, response of the sensor, etc. Indeed, the camera mounted on the roof of the moving vehicle acquires the sky in the presence of clouds and sun rays with high brightness. Therefore, images show brilliant peaks due to the variation in brightness and may present saturated grayscale and, homogeneous and bright local regions (see Fig. 7). If any segmentation algorithm is applied directly to image segmentation without any pre-processing step, the problem of over-segmentation caused by insignificant structures or noise will reach important levels. To deal with this drawback and therefore to extract sky regions with high accuracy, we adopt a common strategy consisting of simplifying the input image with a suitable CI to obtain invariant signatures [11–14]. Then, we try to observe the effects of a certain number of low-level features invariants on the segmentations obtained from SRM algorithm. In this paper, we used 21 CI summarized in Table 1.

3.2. Exponential transform (ET)

The second component of image simplification module is image enhancement. This step consists in improving the interpretability or perception of information in the images to provide better input for the automated image processing steps. This pre-processing step is considered in applications where images are ought to be understood and analyzed like in medical/satellite imaging and natural environment imaging with high dynamic range, such as in our application. In this work, this useful step is used in order to limit the problem of over-segmentation further, keeping all of the regions of interest. There are lots of existing approaches that could be considered for image enhancement. In this work, several efficient algorithms like Peer Group filtering [27], Mean Curvature Motion filtering (based on the diffusion equation) [28], image simplification based on anisotropic smoothing [29], exponential transform (ET), etc., have been tested and evaluated. According to the obtained results, we have opted for exponential transform (ET), which provides the best performances in terms of classification rate. The effect of ET is to approximate the exponential correction factor of grayscale images which maximizes the contrast of the images in the class of exponential intensity mapping functions. In other words, ET enhances detail in high-value regions of the image (bright) while decreasing the dynamic range (defined as $I_{max} - I_{min}$) in low-value regions (dark). Such effect is illustrated in Fig. 3 where it can be seen that the histograms calculated from the image simplified with the ET are likely to contain much more discriminating information than with those obtained on the initial image. ET is given by the following formula (cf. Eq. (1)):

$$\begin{cases} I'_{ij} = exp(\chi/\xi) - 1 + I'_{min} \\ \chi = I_{ij} - I_{min} \\ \xi = (I_{max} - I_{min})/(log(I'_{max} - I'_{min} + 1)) \end{cases} \quad (1)$$

where I_{ij} is the intensity of the pixel at position (i, j), I_{max} and I_{min} are the highest and lowest intensities of the image I, respectively and ξ is a normalization factor for stretching output values between the new lowest I'_{min} and highest I'_{max} intensities of the resultant image I'.

Table 1
The 21 tested colorimetric invariants.

Colorimetric invariants		Acronym	Reference
Three-dimensional spaces	- Affine normalization	- Affine	- [15]
	- Greyworld normalization	- Greyworld	- [16]
	- RGB-rank		- [17]
	- Color Constant Color Indexing	- CCCI	- [17]
	- MaxRGB normalization	- MaxRGB	- [17]
	- Intensity normalization	- Chromaticity	- [18]
	- Comprehensive color normalization	- Comprehensive	- [19]
	- c1c2c3		- [20]
	- m1m2m3		- [20]
	- l1l2l3		- [20]
	- l4l5l6		- [21]
	- c4c5c6		- [21]
	- CrCgCb		- [22]
	- m4m5m6		- [22]
	- A1A2A3		- [22]
	- Standard L_2	- L2	-[22]
	- Maximum-intensity normalization -hsl	- Mintensity	- [23]
Bidimensional space	- CrCb space	- CrCb	- [24]
	- Opposite colors (o1o2)	- o1o2	- [25]
	- Reduced coordinates	- RedCoord	- [26]

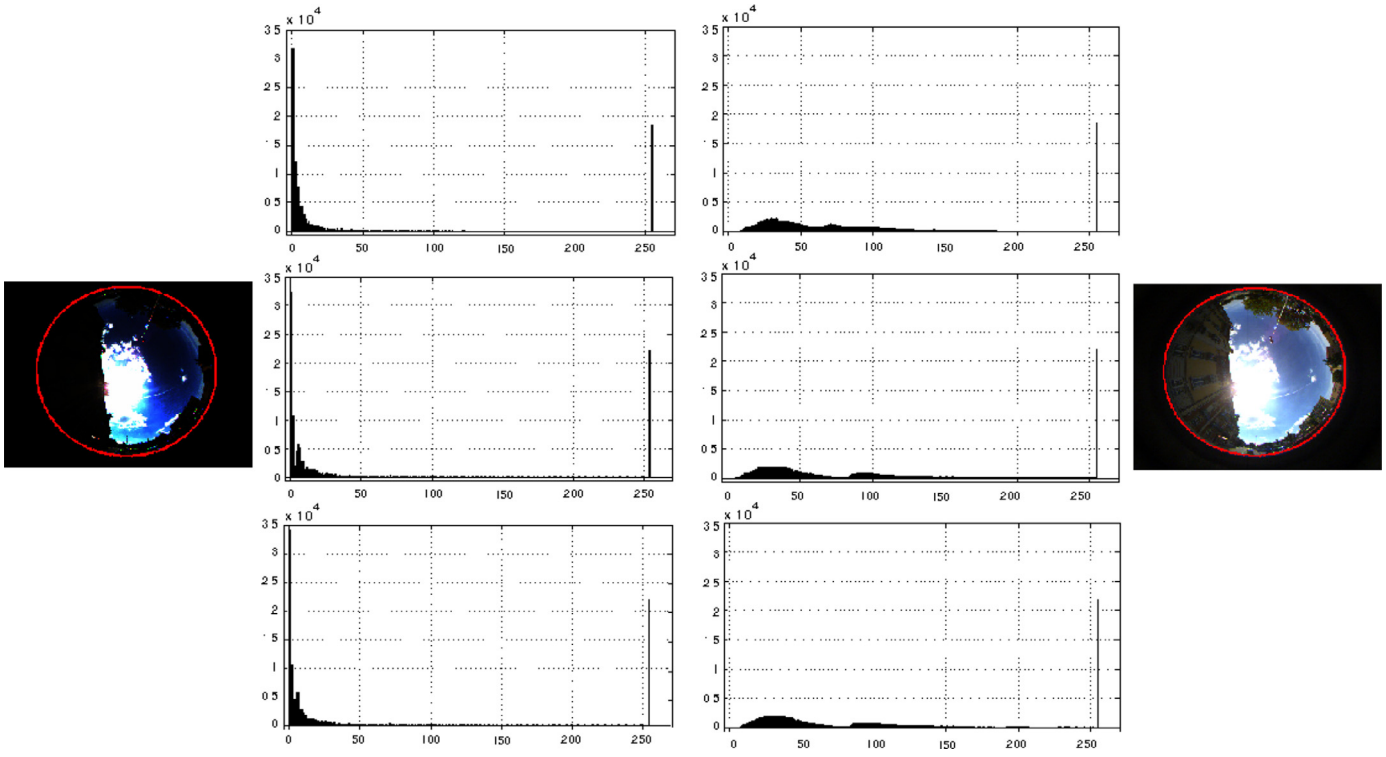

Fig. 3. Histograms of each color channel in the RGB color space (top to bottom red, green and blue channel histograms) obtained without(right) and with exponential transform (left) for the ROI object given by the mask in red color. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Note that, the ET is associated conjointly with the use of color invariance analysis. So, the simplified image, which is the input of the SRM algorithm, is strongly dependent on the CI used. Thus, our analysis consists of determining the optimal couple of CI/ET in the image simplification step. Tests have validated the interest in using this appropriate couple of CI/ET in the segmentation process. Indeed, we will show in the experimental results section that the use of this suitable couple permits to remove the image details, which are not necessary to image segmentation while preserving the contours of the target objects (sky regions).

Fig. 4 highlights the influence of the simplification step by using the optimal couple of CI/ET. One can note that the use of this couple appears to be a good compromise between sufficient attenuation of over-segmentation and proper restitution of the sky-regions. Indeed, the corresponding segmentation is pertinent with good and correct boundary delineation of the desired object and with a low over-segmentation (only 8 regions).

4. Initial segmentation using Statistical Region Merging

The second step of the proposed procedure consists in obtaining a preliminary fisheye image segmentation that correctly extracts all significant regions where boundaries coincide as closely as possible with the significant edges present in the image. Of

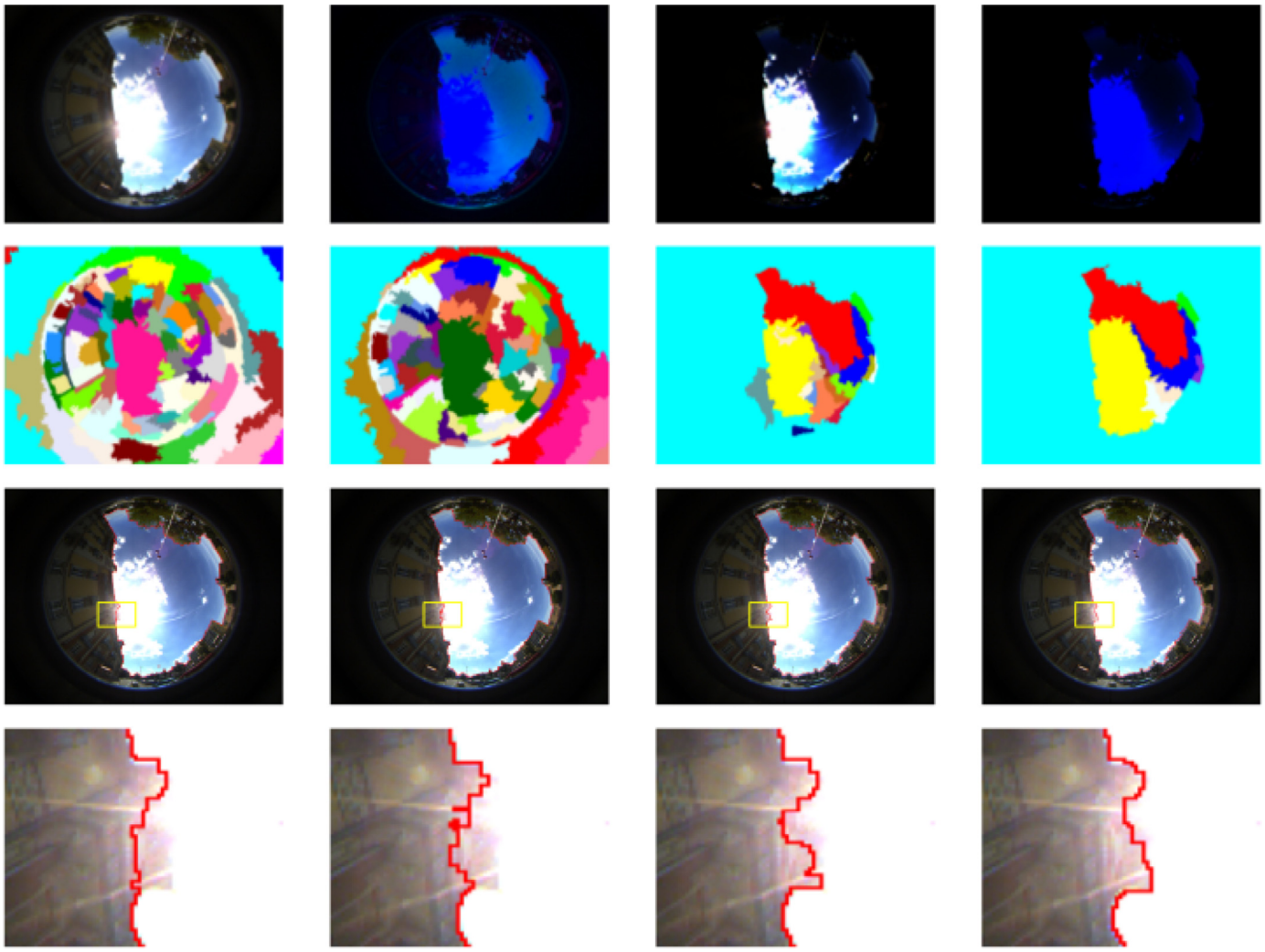

Fig. 4. Influence of simplification step. (first row) Initial images without/with transformation. left to right: original image, image with CI, image with ET and image with CI and ET conjointly; (second row) segmented images with: 81 regions, 66 regions, 20 regions and 8 regions, respectively; (third row) classification results (cf. Section 7); (fourth row) zoom of the rectangle zone in images of the third row.

course, there are many low level segmentation methods in the literature which can achieve that. One can cite Mean shift [30], Jseg unsupervised segmentation algorithm [31], watershed, Turbopixels [32], Statistical Region Merging (SRM) [33], etc. In this paper, we have chosen SRM algorithm to segment the simplified image into homogeneous regions with the same properties (see third second row of Fig. 4). Particular advantages of using this algorithm for dealing with large images are that SRM dispenses dynamical maintenance of region adjacency graph (RAG). It allows defining a hierarchy of partitions and it runs in linear-time by using bucket sorting algorithm while transversing the RAG. In addition, the SRM segmentation method not only considers spectral, shape and scale information, but also has the ability to cope with significant noise corruption and handle occlusions.

5. Region feature extraction

In this stage of our method, we dispose of a segmented image obtained via the SRM algorithm. It is still a challenging problem to accurately extract the object contours from this image because only the segmented regions are calculated and no information estimation on their content, which is necessary for the extraction process, is yet available. Our main goal consists in classifying each segmented region as target object or background. For this purpose, we need first to join the strategy adopted by many authors and which characterizes the regions using suitable descriptors.

It appears from the literature that there are several aspects that could be considered for representing a region such as color, texture, size, shape, and edge [34,35]. These two last features (shape and edge) are unfortunately difficult to describe, as the regions provided by the initial segmentation often change their shape and edge in the images, in particular in fisheye images (non-matrix nature) due to the presence of significant distortions. Identically, region size, although it can be measured simply by computing the number of pixels, it does not allow a unique distinction of objects of interest since they can have different sizes from an image to another or simply they can have the same size as other objects belonging to the background of the image. In this research, we investigate the performance provided by the color and/or texture information. For that, an evaluation of a proposed comprehensive set of 22 local color histograms (rgb color histogram and 21 local color invariance histograms), 10 color local texture histograms and a number of 31 local hybrid descriptors is performed. The descriptors we have chosen for the tests are explained below.

5.1. Local color histograms

Color information, which can be tackled using simply by computing its mean value or its histogram, is an effective parameter to describe statistical information of object color distribution. Note that region histograms are local histograms and they reflect local features in images. Therefore, we exploit color histogram to represent all regions of the segmented image. Recently, Zhu et al. [36] suggested to calculate color histograms using different color space (RGB, HSV, and Opponent color spaces). Six color histograms, obtained by concatenation of the histograms calculated over each channel of the color image, have been proposed. In this paper 22 color histograms are used: (1) RGB color histogram and (2) 21 color invariance histograms.

5.1.1. Local color RGB histograms

A widely used feature in image retrieval and image classification problems is the RGB color histogram, which is as well an important feature for image representation [37,38]. The RGB color histogram of an image is invariant with rotation, translation, and scale [39]; therefore, it is very suitable for color based image classification. In this paper, in order to produce the RGB color histograms, first we had recourse to uniformly quantize each RGB color channel into l levels; afterwards, the color histogram of each region is calculated in the feature space of $z = l \times l \times l$ bins. Some experiments with small numbers of bins have been undertaken, good results have been reached with l=16 bins. It is preferable to compute the histogram from the highest spatial resolution available. However, we did not try to increase this number, because it is computationally too intensive. Given an image I containing N pixels quantified in z=4096 color bins, the RGB color histogram of a segmented region $\mathcal{R}$ is represented as

$$\mathcal{H}ist^{RGB}(\mathcal{R}) = [\mathcal{H}ist^{1}_{\mathcal{R}}, \mathcal{H}ist^{2}_{\mathcal{R}}, \ldots, \mathcal{H}ist^{z}_{\mathcal{R}}] \quad (2)$$

where

$$\mathcal{H}ist^{i}_{\mathcal{R}} = \sqrt{\left(\frac{\sum_{j=1}^{N} p_{i|j}}{\tau}\right)}; \mathrm{j} \in \mathcal{R} \text{ and } 1 \leqslant i \leqslant z. \quad (3)$$

$\mathcal{H}ist^{i}_{\mathcal{R}}$ is the ith normalized histogram bin and τ=card($\mathcal{R}$) is the number of pixels in the region $\mathcal{R}$. $p_{i|j}$ is the conditional probability of the selected jth pixel belonging to the ith color bin. It is expressed as follows:

$$p_{i|j} = \begin{cases} 1, \text{if the } j\text{th pixel is quantized into the } i\text{th color bin} \\ 0, \text{otherwise.} \end{cases} \quad (4)$$

5.1.2. Local color invariance histograms

In order to increase the photometric invariance properties and enhance the discriminative performance, we extend the RGB color histogram to different CI and propose 21 local color invariance histograms that are invariants with different kinds of illumination changes (shadows, brightness, etc.). The 21 incorporated CI ($\mathcal{H}ist^{L_2}$, $\mathcal{H}ist^{Greyworld}$, $\mathcal{H}ist^{c1c2c3}$, $\mathcal{H}ist^{RGB-rank}$, etc.), summarized in Table 1, are calculated by following the same way as for RGB color histogram.

5.2. Color local texture histograms

Texture analysis is an area of intense research. The keen interest in this topic stems from the important role that it plays in many disciplines and related applications: computer assisted diagnosis, remote sensing, surface grading, defect detection and food inspection are just some examples where texture analysis, is, nowadays, a standard. Several general-purpose algorithms and techniques have been developed for texture analysis in the state-of-the-art [40,41].

Table 2
Texture descriptors extracted and used in this work.

Descriptors	Acronym	Reference
Orthogonal combination of local binary pattern	OCLBP	[36]
Local binary patterns	LBP	[44]
Local quinary patterns	LQP	[45]
Local ternary patterns	LTP	[46]
Improved local ternary patterns	ILTP	[47]
Center-symmetric local binary patterns	CS-LBP	[48]
3D Local Binary Patterns	3DLBP	[49]
Local derivative pattern	LDP	[50]
Local Phase Quantisation	LPQ	[51]
Sum and difference histograms	SDH	[52]

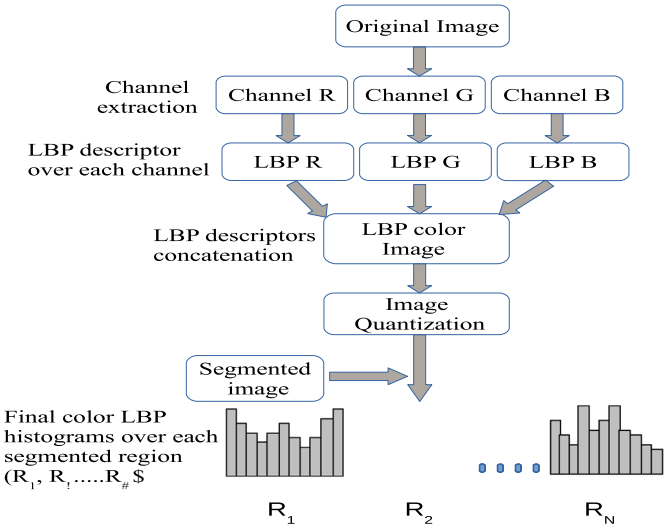

Fig. 5. Calculation of color LBP histogram over each region of the segmented image.

There has been a limited but increasing amount of works that exploit the color aspects of textured images [36,42,43]. However, the selection of suitable texture descriptors depends on the type of objects to be extracted. In this study, we investigate the effect of several existing texture descriptors on the outcome of supervised classification-based sky extraction. For that, a number of 10 texture features calculated for each region of the segmented image are used (cf. Table 2). In order to incorporate color information, we extend these 10 texture descriptors to RGB color space and propose 10 color local texture histograms ($\mathcal{H}ist^{LBP}$, $\mathcal{H}ist^{LTP}$, $\mathcal{H}ist^{LDP}$, etc.). The main idea is to calculate the unichrome texture feature independently over different channels in a RGB color space, concatenate them to get a descriptor color image (for example the LBP color image in Fig. 5) and then the color local texture histogram is calculated for each region of the segmented image by following the same steps as for RGB color histograms, as shown in Fig. 5. This procedure is desirable in the sense that it is required to transform the unichrome texture feature, separately and independently computed from each color-component image, into a more compact, informative and fixed-length representation for more precise classification. Note that, the same procedure is used in order to calculate the color histograms for the others tested texture descriptors.

5.3. Local hybrid histograms

Hybrid descriptors can be obtained by concatenating the feature vectors provided by different descriptors. Given two local color histograms $\mathcal{H}ist(\mathcal{R}) = [\mathcal{H}ist^{1}_{\mathcal{R}}, \mathcal{H}ist^{2}_{\mathcal{R}}, \ldots, \mathcal{H}ist^{z}_{\mathcal{R}}]$ and $\mathcal{H}ist'(\mathcal{R}) = [\mathcal{H}ist^{1'}_{\mathcal{R}}, \mathcal{H}ist^{2'}_{\mathcal{R}}, \ldots, \mathcal{H}ist^{z'}_{\mathcal{R}}]$ of a region $\mathcal{R}$, the corresponding local

hybrid color histogram is mathematically represented as

$$\begin{aligned}\mathcal{H}ist^{hyb}(\mathcal{R}) &= [\mathcal{H}ist_{\mathcal{R}}\ \mathcal{H}ist'_{\mathcal{R}}] \\ &= [\mathcal{H}ist^{1}_{\mathcal{R}}, \mathcal{H}ist^{2}_{\mathcal{R}}, \ldots, \mathcal{H}ist^{z}_{\mathcal{R}}, \mathcal{H}ist^{1'}_{\mathcal{R}}, \\ &\quad \mathcal{H}ist^{2'}_{\mathcal{R}}, \ldots, \mathcal{H}ist^{z'}_{\mathcal{R}}]\end{aligned} \tag{5}$$

where $\mathcal{H}ist^{i}_{\mathcal{R}}$ is *ith* histogram bins and the dimension of the obtained $\mathcal{H}ist^{hyb}(\mathcal{R})$ will be $(2 \times z)$ sized; z=4096.

In this study, we propose to concatenate color RGB histogram on the one hand with color invariance histograms leading to 21 hybrid color invariance histograms ($\mathcal{H}ist^{RGB}_{c1c2c3}$, $\mathcal{H}ist^{RGB}_{Affine}$, $\mathcal{H}ist^{RGB}_{A1A2A3}$, etc.) and on the other hand with texture histograms leading to 10 hybrid color texture histograms ($\mathcal{H}ist^{RGB}_{LBP}$, $\mathcal{H}ist^{RGB}_{LTP}$, $\mathcal{H}ist^{RGB}_{ILTP}$, etc.). While this can enrich the discrimination capacity of the resulting descriptor, it has the disadvantage that the dimensionality of the resulting feature vector could be very high, increasing thus the computation time.

6. Maximal similarity based region classification

Now that we have defined the feature adopted for characterizing the regions of the segmented image $\mathcal{M}_{\mathbf{SRM}}$ via SRM algorithm, the key issue is to determine similarity between the regions of $\mathcal{M}_{\mathbf{SRM}}$ and those of two learning databases $\mathcal{B}_{\mathbf{obj}}$ and $\mathcal{B}_{\mathbf{back}}$ that are constructed respectively with m distinctive textures of objects to be extracted (sky regions in our case) and n distinctive textures of other objects such as building, road, tree, etc. For that, we need to define a similarity measure rule $\varrho(\mathcal{R}, \mathcal{Q})$ between two regions $\mathcal{R}$ and $\mathcal{Q}$ basing on their descriptors. The most similarity measures commonly used are based on vector space model, i.e. taking image region features as points in the vector space, through the calculation of close degree of two points to measure the similarities between the image region features. There are some well-known goodness-of-fit statistical metrics such as Minkowski measure, histogram intersection method [53], second type distance [54], Bhattacharyya coefficient [55,56], log-likelihood ratio statistic [57], etc. For regions $\mathcal{R}$ and $\mathcal{Q}$, using the notation $\varrho(\mathcal{R}, \mathcal{Q})$ representing the similarity between them, the larger ϱ is, the higher similarity between regions $\mathcal{R}$ and $\mathcal{Q}$ we will get. Denote by $\mathcal{H}ist^{i}_{\mathcal{R}}$ the normalized histogram of a region $\mathcal{R}$, the superscript i represents its ith element. $z = l \times l \times l = 4096$ represents the feature space.

In this work, we adopted Bhattacharyya coefficient (cf. Eq. (6)), which represents the cosine of angle between the unit vectors representing the two regions to be compared:

$$\left(\sqrt{\mathcal{H}ist^{1}_{\mathcal{R}}}, \ldots, \sqrt{\mathcal{H}ist^{z}_{\mathcal{R}}}\right)^{T}$$

and

$$\left(\sqrt{\mathcal{H}ist^{1}_{\mathcal{Q}}}, \ldots, \sqrt{\mathcal{H}ist^{z}_{\mathcal{Q}}}\right)^{T}$$

This choice is due to the ability of this measure to model closely the similarity value of the two formed vectors. The higher the Bhattacharyya coefficient between regions $\mathcal{R}$ and $\mathcal{Q}$ is, the higher the similarity between them is. That is to say their histograms are very similar and the angle between the two histogram vectors is very small. Certainly, two similar histograms do not necessarily involve that the two corresponding regions are perceptually similar. Nevertheless, coupling with the proposed $\mathcal{MSRC}$ algorithm summarized as in Algorithm 1, Bhattacharyya similarity works well in the proposed approach. It might be worth to mention that a histogram is a global descriptor of a local region and it is robust to noise and small variations. Given that the Bhattacharyya coefficient is the inner product of two histogram vectors, this coefficient is thus robust to noise and small variations too.

$$\varrho(\mathcal{R}, \mathcal{Q}) = \sum_{i=1}^{z} \sqrt{\mathcal{H}ist^{i}_{\mathcal{R}}.\mathcal{H}ist^{i}_{\mathcal{Q}}} \tag{6}$$

We may further notice that the proposed $\mathcal{MSRC}$ algorithm can be seen as a variant of k-Nearest Neighbor (KNN) method where both methods are conceptually quite simple. However, in KNN procedure, an object is classified according to the classes of the k closest objects, i.e. it is assigned to the class to which belong the majority of k neighboring objects. This assignment is performed without taking into account any relation between objects inside a given class. However, in the context of regions classification, given that two similar color histograms do not necessarily involve that the two corresponding regions are perceptually similar, considering only vote rule to classify regions seems to be insufficient. Therefore, unlike KNN, the key decision rule of the proposed method consists in assigning an unknown region R to the class C_n, if the average of the K first high similarity measures (in terms of the Bhattacharyya measure) calculated between the region R and the regions of the learning database corresponding to the class C_n is maximal, i.e,

$$C(\mathcal{R}) = \arg\max_{C_n \in C} \frac{1}{K} \sum_{i=1}^{K} \varrho(\mathcal{R}, \mathcal{Q}_i),\ \mathcal{Q}_i \in B_n \subset D \tag{7}$$

where D=$\{B_1, B_2, \ldots, B_l\}$ is a set of learning databases corresponding to the classes C=$\{C_1, C_2, \ldots, C_l\}$, $\mathcal{R}$ is a query, and ϱ is a similarity measure rule.

The $\mathcal{MSRC}$ algorithm becomes equivalent to KNN method when $K = 1$ and if the same measure is used to calculate distance between regions. As for KNN method, the optimal value of K is searched by means of cross validation procedures, i.e. by testing a set of values for K (e.g. from 1 to 10); then, the value giving the lowest classification error in cross-validation can be selected as the optimal one.

Algorithm 1 Maximal similarity based region classification.

Require: $I \leftarrow$ input image.
$\mathcal{B}_{\mathbf{obj}} \leftarrow$ learning database of objects of interest (sky regions).
$\mathcal{B}_{\mathbf{back}} \leftarrow$ learning database of background (building, road, tree, etc.).
Step 1 : **Image Simplification and Segmentation**
Simplify I through the optimal couple of CI/ET to obtain the simplified image J.
Segment J into regions through SRM algorithm in order to obtain the set $\mathcal{M}_{\mathbf{SRM}}$ of segmented regions.
Step 2 : **Region feature extraction**
Calculate the histogram feature of all regions of $\mathcal{M}_{\mathbf{SRM}}$, $\mathcal{B}_{\mathbf{obj}}$ and $\mathcal{B}_{\mathbf{back}}$.
Step 3 : **Region Classification**
Substep 3.1 : **Distance Calculation**
for each candidate region $\mathcal{R} \in \mathcal{M}_{\mathbf{SRM}}$ **do**
  Calculate the similarity vector $V^{\mathcal{R}}_{obj} = \{\varrho(\mathcal{R}, \mathcal{Q}_i); (\mathcal{Q}_i)_{i=1..m} \in \mathcal{B}_{\mathbf{obj}}\}$ between $\mathcal{R}$ and $\mathcal{B}_{\mathbf{obj}}$. $\varrho(\mathcal{R}, \mathcal{Q}_i)$ is the similarity between the region $\mathcal{R}$ and the region $\mathcal{Q}_i \in \mathcal{B}_{\mathbf{obj}}$.
  Calculate the similarity vector $V^{\mathcal{R}}_{back} = \{\varrho(\mathcal{R}, \mathcal{Q}_j); (\mathcal{Q}_j)_{j=1..n} \in \mathcal{B}_{\mathbf{back}}\}$ between $\mathcal{R}$ and $\mathcal{B}_{\mathbf{back}}$. $\varrho(\mathcal{R}, \mathcal{Q}_j)$ is the similarity between the region $\mathcal{R}$ and the region $\mathcal{Q}_j \in \mathcal{B}_{\mathbf{back}}$.
  Get the order of $V^{\mathcal{R}}_{obj}$ and $V^{\mathcal{R}}_{back}$ by decreasing sorting;
  Calculate $\mu^{\mathcal{R}}_{obj} = \frac{\sum_{i=1}^{K} \varrho(\mathcal{R}, \mathcal{Q}_i)}{K}$, $K \leq m$, the mean of the K first elements of $V^{\mathcal{R}}_{obj}$.
  Calculate $\mu^{\mathcal{R}}_{back} = \frac{\sum_{j=1}^{K} \varrho(\mathcal{R}, \mathcal{Q}_j)}{K}$, $K \leq n$, the mean of the K first elements of $V^{\mathcal{R}}_{back}$.
Substep 3.2 : **Decision rule**
  if $(\mu^{\mathcal{R}}_{obj} \geq \mu^{\mathcal{R}}_{back})$ **then**
    The region $\mathcal{R}$ has the maximal similarity with $\mathcal{B}_{\mathbf{obj}}$, it is then classified as a part of sky regions.
  else
    The region $\mathcal{R}$ has the maximal similarity with $\mathcal{B}_{\mathbf{back}}$, it is then classified as a part of background.
  end if
end for
return The final segmentation map.

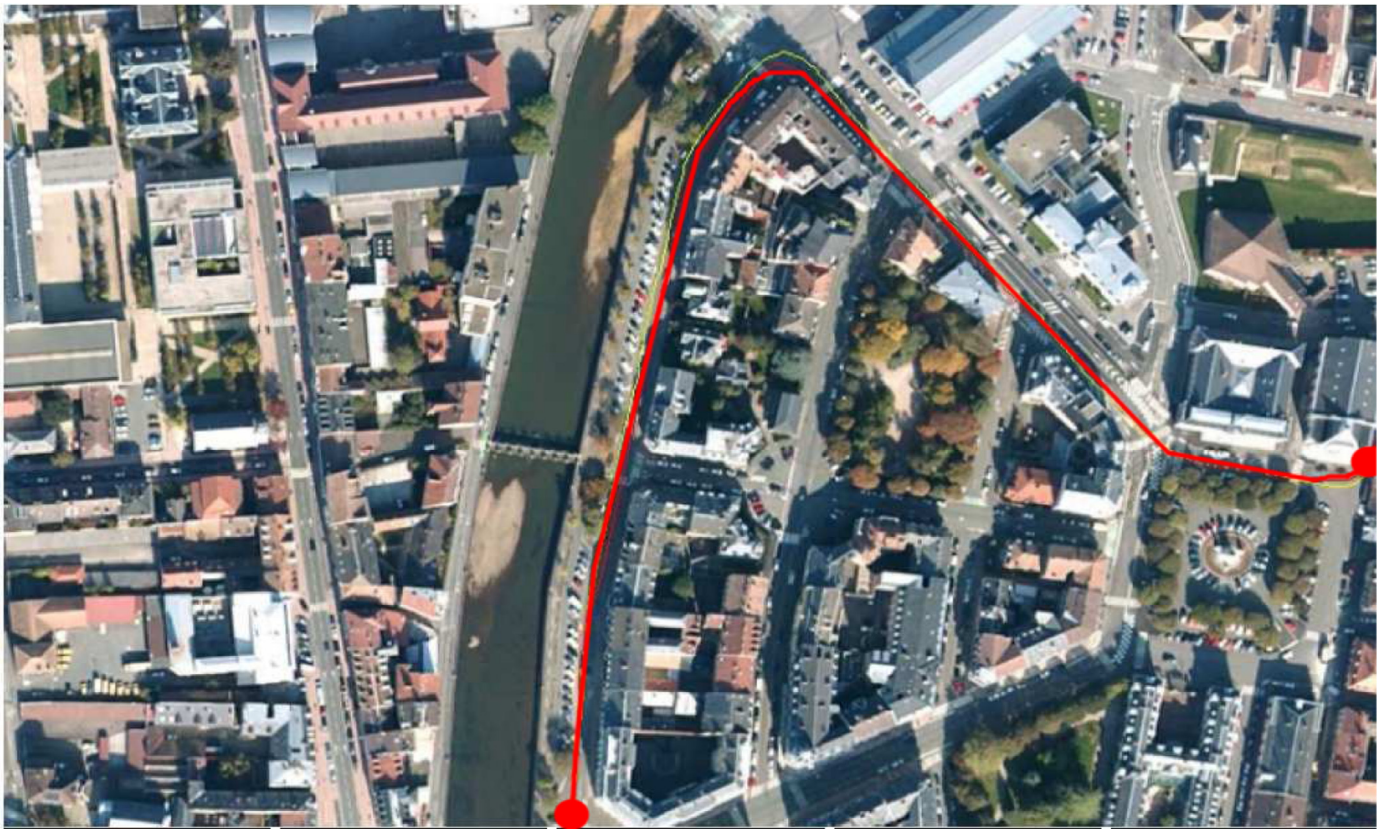

Fig. 6. The global trajectory during the database acquisition, within the framework of the CAPLOC project.

7. Results and discussion

An extensive experimental study was carried out to investigate the effectiveness of the proposed $\mathcal{MSRC}$ method and the proposed groups of local color invariance, color local texture and local hybrid image region descriptors for classification of regions of interest in fisheye images. A comparison with the method in [9] was performed to show the improvement that our method provides. In order to prove the performance of our proposed algorithm, another comparison with the algorithm in [10] was also performed. The following sections describe in detail the experimental results of the different steps of the proposed approach, namely: (1) Choice of the appropriate couple of CI/ET; (2) $\mathcal{MSRC}$ and descriptors performance; and (3) Comparative evaluation and discussion, as well as the image database and evaluation metrics used. To better understand and visualize the performance of each step of our proposed approach, we have chosen to show the results in both qualitative and quantitative ways. We point out that for better readability, we have only presented the results for best tested descriptors.

7.1. Dataset

The image database acquired in the framework of the CAPLOC project, was used to validate the proposed sky detection/extraction procedure. The database was acquired in June 2010 in Belfort situated on the north-eastern of France. It has heterogeneous data and varying complex scenarios (urban canyon, vegetation abundance, overexposure, brightness changes, etc.). Fig. 6 illustrates the global trajectory traveled by the experimental vehicle during the database acquisition. The image database contains 150 images exhibiting the mentioned various complex situations. Fig. 7 Illustrates six images of the database.

7.2. Evaluation metrics

In order to get a quantitative evaluation, we use ground-truth fisheye maps. The manually delineated sky regions were used as a reference sky set to assess the automated sky-extraction accuracy. The extracted results and ground-truth ones are compared pixel-by-pixel. Each pixel in the images is categorized as one of the four following types.

1. True positive (TP): Both manual and automated methods label the pixel belonging to sky region.
2. True negative (TN): Both manual and automated methods label the pixel belonging to background.
3. False positive (FP): The automated method incorrectly labels the pixel as belonging to sky region.
4. False negative (FN): The automated method incorrectly labels the pixel as belonging to background.

From these measures it is straightforward to compute the following scores associated with the sky regions in the test images: recall, precision, F1 measure, accuracy, and Matthews Correlation Coefficient (MCC) (cf. Eqs. 8–12). The MCC is in essence a correlation coefficient between the observed and predicted binary classifications; it returns a value between −1 and +1. A value of +1 represents a perfect prediction, 0 no better than random prediction and −1 indicates total disagreement between prediction and observation.

$$\text{recall} = \frac{TP}{TP+FN} \times 100 \tag{8}$$

$$\text{precision} = \frac{TP}{TP+FP} \times 100 \tag{9}$$

$$F1 = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}} \tag{10}$$

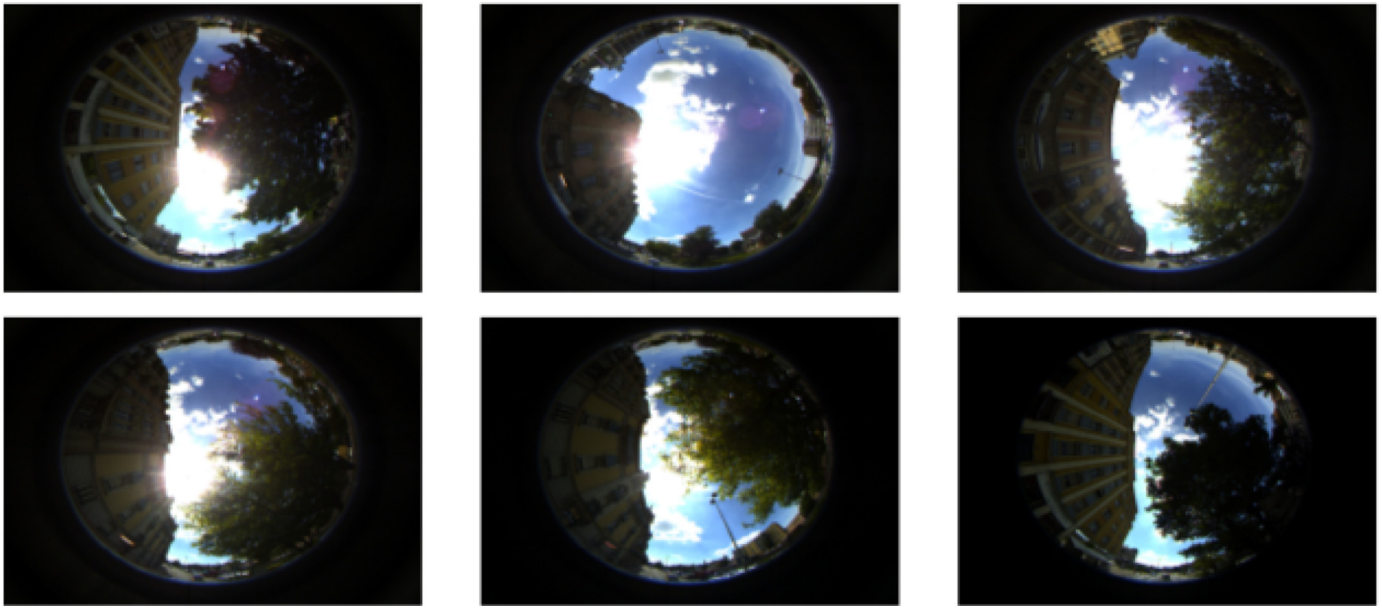

Fig. 7. Six images of the database (acquired in the framework of the CAPLOC project) with varying complexity (overexposure, brightness changes, presence of clouds and strong illumination, etc.).

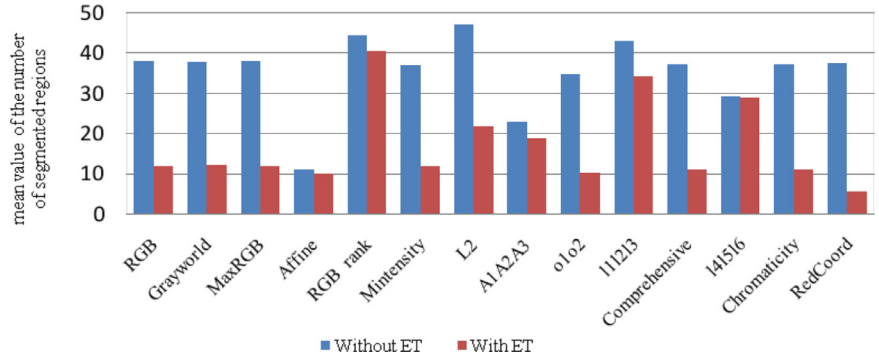

Fig. 8. Attenuation of over-segmentation problem according to the use of the couple CI/ET.

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \tag{11}$$

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \tag{12}$$

7.3. Choice of the appropriate couple of colorimetric invariant/exponential transform

As indicated in Section 3, we propose to simplify the initial image using an appropriate couple of CI/ET transform where the effect is to limit illumination changes and thus reduce the over-segmentation problem. Fig. 8 illustrates the attenuation of over-segmentation problem according to the use of different couples of CI/ET. We point out that for better readability, we have only presented 14 best invariants among the 21 tested. Each bar represents the mean value of the number of segmented regions calculated on all images of the test database according to the couple CI/ET tested. One can see that when using the CI and the ET conjointly, the number of regions is considerably reduced, and this, for the most tested CI. Considering the region reduction rate, we can rank the best couples as follows: rc/ET (reduction rate of 84.65%), Affine/ET (reduction rate of 73.46%), o1o2/ET (reduction rate of 72.62%), Chromaticity/ET (reduction rate of 70.3%), Comprehensive/ET (reduction rate of 70.27%), etc. Nevertheless, if we consider the evaluation of classification results, we notice that the couple Affine/ET appears to be a good compromise between sufficient attenuation of over-segmentation and proper restitution of sky regions (cf. Table 3). Indeed, using the couple rc/ET considerably reduces the over-segmentation of the images (providing in

Table 3
Performance of the $\mathcal{MSRC}$ method based on the RGB color histogram without/with ET.

Tested CI	Recall	Precision	Fmeasure	Accuracy	MCC
RGB	99.29	95.93	97.56	99.00	0.97
RGB+ET	98.43	97.41	97.90	99.17	0.97
RedCoord	87.08	86.19	86.28	94.29	0.83
RedCoord+ET	85.55	96.75	90.03	96.34	0.89
MaxRGB	99.29	95.93	97.56	99.00	0.97
MaxRGB+ET	98.43	97.41	97.90	99.17	0.97
RGB-rank	91.03	93.61	92.19	96.94	0.90
RGB-rank+ET	98.28	95.72	96.94	98.77	0.96
o1o2	99.38	96.36	97.82	99.12	0.97
o1o2+ET	98.54	97.44	97.97	99.20	0.97
Affine	97.18	97.73	97.43	98.98	0.97
Affine+ET	99.36	96.42	97.85	99.13	0.97

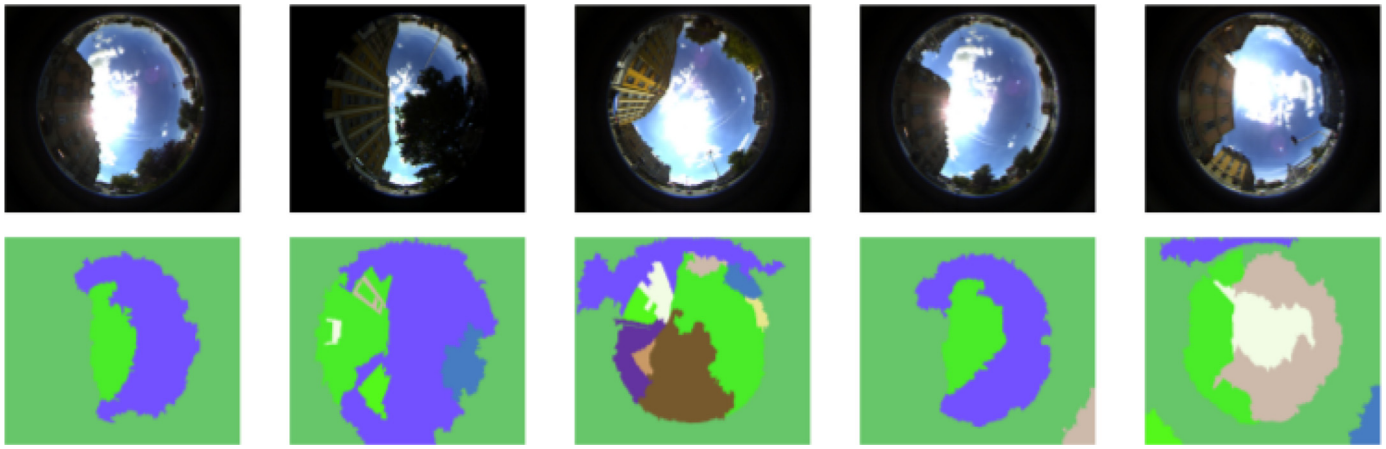

Fig. 9. Under-segmentation problem: presence of mixed regions that are provided by the couple rc/ET.

Table 4
Classification results according to the local color invariance histograms used.

Descriptors	Recall	Precision	Fmeasure	Accuracy	MCC
RGB	99.36	96.42	97.85	99.13	0.97
A1A2A3	99.14	96.87	97.98	99.17	0.97
Affine	98.70	96.66	97.64	99.05	0.97
RedCoord	99.14	96.80	97.94	99.16	0.97
Greyworld	99.11	97.25	98.15	99.26	0.98
l1l2l3	98.77	96.56	97.63	99.02	0.97
l4l5l6	99.03	96.14	97.55	98.97	0.97
Maximum intensity	99.11	97.25	98.15	99.26	0.98
MaxRGB	99.11	97.25	98.15	99.26	0.98
L2	98.54	97.36	97.93	99.18	0.97
hsl	99.03	96.06	97.45	99.03	0.97
o1o2	98.11	97.26	97.66	99.10	0.97
m1m2m3	98.54	96.38	97.37	99.00	0.97
c4c5c6	99.15	95.38	97.20	98.82	0.96

Table 5
Classification results according to the color local texture histograms used.

Descriptors	Recall	Precision	Fmeasure	Accuracy	MCC
LBP	95.29	97.09	96.14	98.43	0.95
OCLBP	99.14	96.87	97.97	99.17	0.97
LQP	85.55	96.75	90.03	96.34	0.89
ILTP	91.86	97.31	94.44	97.81	0.93
LPQ	72.75	96.16	81.15	93.42	0.80
LTP	90.44	95.88	92.22	97.29	0.91
CSLBP	94.10	92.24	92.76	96.92	0.91
3DLBP	72.03	96.99	81.36	93.33	0.80
LDP	87.27	94.90	90.56	96.11	0.89
SDH	99.15	96.36	97.70	99.07	0.97

Table 6
Classification results according to the color hybrid histograms used.

Descriptors	Recall	Precision	Fmeasure	Accuracy	MCC
RGB ∪ LBP	99.13	97.15	98.11	99.24	0.98
RGB ∪ OCLBP	99.14	97.05	98.07	99.22	0.98
RGB ∪ LQP	99.13	96.88	97.97	99.18	0.97
RGB ∪ ILTP	99.13	97.15	98.11	99.24	0.98
RGB ∪ LPQ	99.13	96.88	97.97	99.18	0.97
RGB ∪ LTP	97.24	97.35	97.27	98.89	0.97
RGB ∪ CSLBP	99.13	96.88	97.97	99.18	0.97
RGB ∪ 3DLBP	99.13	96.88	97.97	99.18	0.97
RGB ∪ SDH	99.14	96.61	97.83	99.13	0.97
RGB ∪ Greyworld	99.13	97.23	98.13	99.24	0.98
RGB ∪ c4c5c6	99.36	96.72	98.00	99.20	0.98
RGB ∪ l1l2l3	99.36	96.42	97.85	99.13	0.97
RGB ∪ l4l5l6	99.41	96.42	97.86	99.15	0.97
RGB ∪ hsl	98.54	97.15	97.82	99.13	0.97
RGB ∪ A1A2A3	99.12	96.92	98.03	99.15	0.97
RGB ∪ m1m2m3	99.40	96.36	97.83	99.13	0.97
RGB ∪ o1o2	99.34	96.39	97.82	99.12	0.97

average 5.8467 regions per image) but the classification results show that it is not suitable for obtaining the preliminary segmentation results. The underlying reason of this shortcoming is that the couple rc/ET based preliminary segmentation leads to segmented images with under-segmentation, i.e. presence of mixed segmented regions (i.e., presence of regions that contain pixels belonging to different classes) in the corresponding segmentation results (cf. Fig. 9). Consequently, the pixels misclassification problem will emerge, i.e. the classifier will affect the same label to the whole pixels belonging to any mixed segmented region leading to some errors independently to the descriptor and the classifier being used. Even if the couple Affine/ET leads to a number of segmented regions, which is relatively higher (in average 10,43 regions per image) in comparison with the couple rc/ET, it provides good classification results where the produced classified images are most close the corresponding ground truth (judged by the evaluation results given in Table 3). In conclusion, we opted to use Affine normalization/Exponential transform as an appropriate couple for image simplification purposes.

7.4. $\mathcal{MSRC}$ and descriptors performance

In this section, we study using the evaluation measures previously introduced, the performance of the proposed $\mathcal{MSRC}$ method and different descriptors to classify the segmented regions and thus extract those corresponding to sky regions. Our goal is to choose the appropriate descriptor which with the $\mathcal{MSRC}$ method is able to correctly classify all regions of the segmented image into sky and non sky regions. In Tables 4–6, we report experimental results obtained using local color invariance histograms, color local texture histograms, color hybrid histograms and color texture hybrid histograms.

Table 4 summarizes the average (on the dataset) of the performance indicators recall, precision, F1 measure, Accuracy and MCC according to the $\mathcal{MSRC}$ procedure and the local color RGB and invariance histograms used. As can be seen, all of the tested color descriptors allows obtaining good classification results with a small increase when using certain local color invariance histograms. It is worthy to notice that the local color invariance histograms Greyworld, Maximum intensity and MaxRGB are the most promising because they give the maximum classification rates compared to the other tested descriptors.

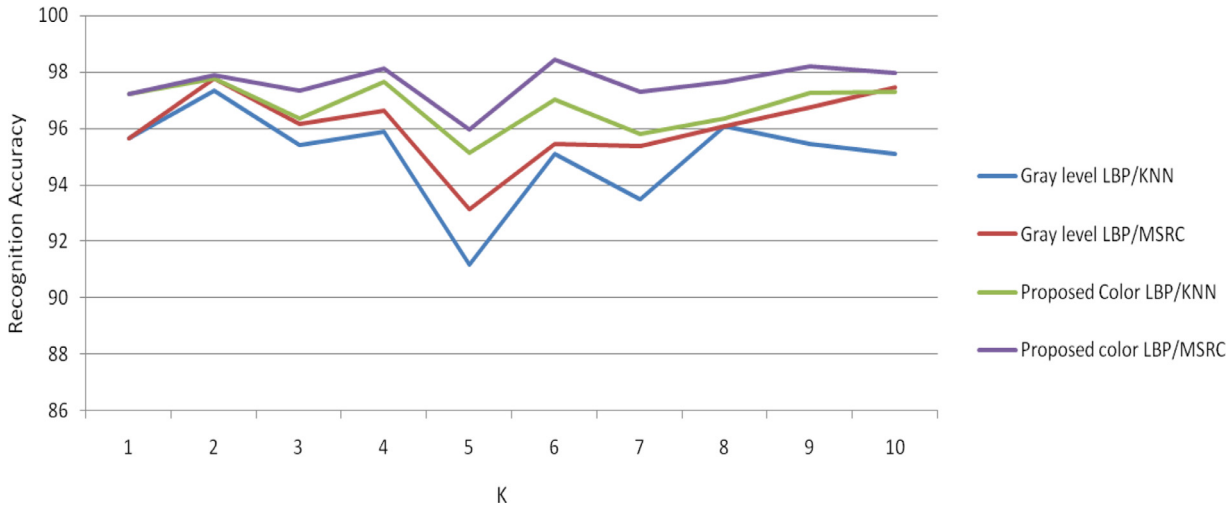

Fig. 10. Gain between the proposed color LBP and MSRC method.

Table 5 is similar to Table 4, however this time the descriptors used are color local texture histograms. One can notice that the SDH and OCLBP descriptors allowed obtaining the best classification rates compared to the other tested descriptors. Table 6 summarizes the performance of the used hybrid descriptors. The analysis of the classification rates leads us to weight the influence of the use of hybrid descriptors. Indeed, when used in a hybrid form with rgb color histogram, all tested color local texture histograms give good results as they permit to increase the classification rate. This increase also concerns certain color invariance histograms while for others the results remain relatively stable.

We may further notice that, the proposed local color invariance descriptors demonstrate their effectiveness as they show superior performances than their texture or hybrid counterparts. The underlying reason could be that the acquired images contain more varieties of illumination changes and do not show significant textural properties. This also explains why normalization intensity based color invariance descriptors perform the best among the others color invariance descriptors, since they possess the strongest invariance properties (invariant to light color change and to illumination of the scene).

In the following, we show the gain between the new local descriptors and MSRC method. For simplicity, we considered as reference the classical LBP (gray level) descriptor and KNN (K-Nearest Neighbors) method as its principle is quite similar to the one used in MSRC. To perform fair evaluation KNN is applied with the same distance as the one used in MSRC (Bhattacharyya coefficient-based distance). Again for simplicity, we considered for this evaluation the proposed color LBP descriptor. Fig. 10 shows the recognition accuracy for the combinations classic LBP/KNN, classic LBP/MSRC, proposed color LBP/KNN and proposed color LBP/MSRC, for different values of K. One can see that both proposed color LBP descriptor (proposed color LBP/KNN) and MSRC method (classic LBP/MSRC) increased the performance, with relatively significant improvement for the proposed color LBP for K varying in [4,6]. The figure confirms once again that the combination proposed color LBP/MSRC method provides the best performance by outperforming all the others.

7.5. Comparative evaluation and discussion

As indicated in Section 1, the study of this paper followed the work initiated within the CAPLOC project [9] where one of the objectives is to classify images into sky and non-sky classes. In this project, the authors have compared the performance of different well known clustering algorithms often used for image classification purposes. Several unsupervised (Fisher, KMlocal, Fuzzy-Cmeans, SRM) and supervised (Bayes, KNN, SVM) classifiers are tested. Our method was also compared to the algorithm described in [10] which is based on pixel classification formulated as a minimization task. It is obvious to note that in the evaluation phase, for meaningful comparison, each algorithm (pre-processing, classification and post-processing) was tested over many possible combinations of input parameters.

Table 7 illustrates the best classification rates for each tested method. It can be seen that the proposed framework behaves better than the popular tested classifiers in [9] and the algorithm proposed in [10]. Indeed, the proposed approach maximizes the classification rate for all the used evaluation metrics. For example, if we consider the MCC measure, our method performs clearly better and gives the best classification rate (98% using the local color invariance histograms as descriptors) compared to other supervised and unsupervised tested methods (96% with ALE, 95% with Fisher, etc.). Thus, the proposed $\mathcal{MSRC}$ algorithm permits to increase the MCC measure with 2%. If we consider the accuracy measure, the proposed approach gives the classification rate of 99.26% vs 97,71% with Fisher vs. 97.67% with KNN, vs. 97.59% with ALE, etc. This shows that our method allows to increase the accuracy with 1,55%.

Fig. 11 illustrates an example of visual comparison of the region classification results. From left to right: acquired image, classified image into two classes (sky and non-sky) obtained with the best classifier defined in [9], classified image obtained with the method in [10] and classification result obtained with the proposed $\mathcal{MSRC}$ method applied with the optimal couple CI/ET and with the defined best local color invariance histograms. Basing on this visual evaluation, one can state that the developed approach demonstrates excellent accuracy in terms of sky boundary extraction. Indeed, for all of the images in the first column of Fig. 11, the produced classification results (cf. fourth column of Fig. 11) agree most closely with the corresponding ground truth where the majority of the sky regions are detected with good boundary delineation. Note that the results obtained by the method proposed in [9], which is based on pixel-by-pixel classification that tries to extract high level information without any knowledge of how pixels should be grouped, is often accompanied with false positives. Indeed, from the second column of Fig. 11, which illustrates classification results obtained with the best classifier defined in [9], one can see clearly that, due to radiometric similarity between some

Table 7
Classification results according to the classifiers used in the CAPLOC project [9], according to the method introduced in [10] and according to our proposed method.

Algorithm	Ref.	Optimized parameters	Recall	Precision	Fmeasure	Accuracy	MCC
Fuzzy-Cmeans		F = 2	99.34	93.49	96.27	96.89	0.94
Fisher		Blue	99.49	95.08	97.19	97.71	0.95
SVM		–	98.96	91.08	94.79	95.53	0.91
Bayes	[9]	–	99.26	94.31	96.68	97.27	0.94
KNN		K = 6	98.07	95.71	96.76	97.67	0.95
SRM		–	98.96	91.08	94.79	95.53	0.91
KMlocal		Loyds	99.37	93.69	96.40	97.03	0.94
ALE	[10]	Default	98.60	96.04	97.29	97.59	0.96
Our approach	This paper	K=6 & Greyworld	99.11	97.25	98.15	99.26	0.98

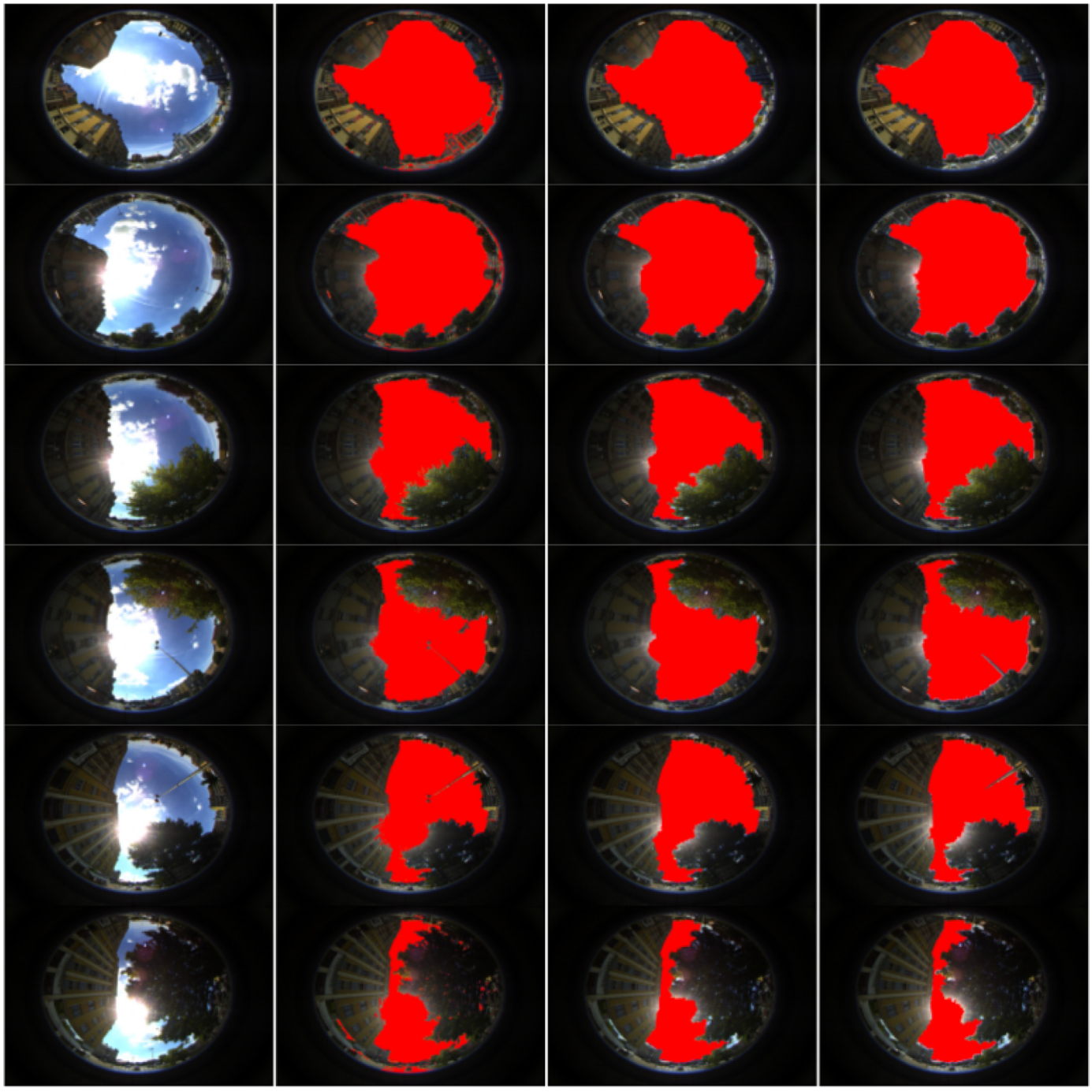

Fig. 11. Visual comparison of region classification results. From left to right: acquired image, classified image into two classes (sky and non-sky) obtained by the best classifier defined in [9], classified image obtained with the method in [10] and classification result obtained by the proposed $\mathcal{MSRC}$ approach.

pixels belonging to building and sky regions, several not-sky parts are classified as true positive (sky region). In addition, some parts of building or tree regions, which are exposed to the sun, are lost and classified as true positive (sky region). The method in [10], which is based also on pixel classification, suffers from the same problem, as illustrated in the third column of Fig. 11 where some parts of building, floor lamp or tree regions are detected as true positive (sky region).

This visual evaluation confirms hence the conclusion given from the analysis of the results of Table 7. The outperformance of the proposed method is possible thanks to its two main steps (new regions representation and MSRC based regions classification), as demonstrated through the analysis of the results reported in Fig. 10.

8. Conclusion

This paper presented a region based classification method for line of sight extraction using fisheye images, in the framework of enhancing GNSS based localization. The proposed method starts by image simplification using an appropriate couple of colorimetric invariant and exponential transform.

Then, we have introduced several new local color descriptors for image region description based on color invariance and texture features. The proposed descriptors incorporate color information to increase their discriminative power, and also to enhance their photometric invariance properties over various illumination changes.

Finally, from regions obtained with the SRM algorithm, a Maximal Similarity Based Region Classification using Bhattacharyya coefficient-based distance has been proposed in order to classify the characterized regions into two classes (sky and non-sky regions). Comparison of our method with the results obtained with the state-of-the-art approaches showed that the proposed method outperforms all the tested methods.

Future works will focus on processing images acquired all day and in different meteorological situations in order to study the ability of the proposed approach to deal with hard scenarios. We envision also night situation by using infrared imaging. Furthermore, we plan to extend the analysis by considering more color spaces and texture descriptors. Another perspective is to extend the evaluation of the proposed framework to other data sets related to applications dealing with automatic objects recognition.

References

[1] B.S. Pervan, D.G. Lawrence, C.E. Cohen, B.W. Parkinson, Parity space methods for autonomous fault detection and exclusion using GPS carrier phase, Proceedings of the Symposium on Position Location and Navigation, 1996, pp. 649–656.

[2] J.-H. Wang, Y. Gao, High-sensitivity GPS data classification based on signal degradation conditions, IEEE Trans. Veh. Technol. 56 (2) (2007) 566–574.

[3] M. Lentmaier, B. Krach, P. Robertson, Bayesian time delay estimation of GNSS signals in dynamic multipath environments, Int. J. Navig. Obs. 11 (2008) 372651. Hindawi Publishing Corporation.

[4] O. Koch, S. Teller, Wide-area egomotion estimation from known 3D structure, in: Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR), 2007.

[5] J.-I. Meguro, T. Murata, J.-I. Takiguchi, Y. Amano, T. Hashizume, GPS multipath mitigation for urban area using omnidirectional infrared camera, IEEE Trans. Intell. Transp. Syst. 10 (1) (2009) 22–30.

[6] S. Ramalingam, S. Bouaziz, P. Sturm, M. Brand, Skyline2gps: localization in urban canyons using omni-skylines, in: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2010.

[7] P.D. Groves, Shadow matching: a new GNSS positioning technique for urban canyons, J. Navig. 64 (2011) 417–430.

[8] A. Bourdeau, M. Sahmoudi, J.-Y. Tourneret, Tight integration of GNSS and a 3D city model for robust positioning in urban canyons, ION GNSS, Nashville, TN, 2012.

[9] J. Marais, C. Meurie, D. Attia, Y. Ruichek, A. Flancquart, Toward accurate localization in guided transport: combining GNSS data and imaging information, Trans. Res. Part C Emerg. Technol. 43 (2013) 188–197.

[10] O. Saurer, G. Baat, K. Köser, L. Ladicky, M. Pollefeys, Image based geo-localization in the Alps, Int. J. Comput. Vis. 116 (3) (2015) 213–225.

[11] T. Cong, L. Khoudour, C. Achard, C. Meurie, O. Lezoray, People re-identification by spectral classification of silhouettes, Signal Process. 90 (2010) 2362–2374.

[12] Y. El merabet, C. Meurie, Y. Ruichek, A. Sbihi, R. Touahni, Orthophotoplan segmentation and colorimetric invariants for roof detection, in: Image Analysis and Processing–ICIAP, Springer, Berlin Heidelberg, 2011, pp. 394–403.

[13] Y. El merabet, C. Meurie, Y. Ruichek, A. Sbihi, R. Touahni, Segmentation d'images aériennes par coopération LPE-régions et LPE-contours, Application á la caractérisation de toitures. Revue Francaise de Photogrammetrie et de Teledetection 206 (2014) 29–44.

[14] Y. El merabet, C. Meurie, Y. Ruichek, A. Sbihi, R. Touahni, Building roof segmentation from aerial images using a line and region-based watershed segmentation technique, Sensors 15 (2015) 3172–3203.

[15] A. Fusiello, E. Trucco, T. Tommasini et, V. Roberto, Improving feature tracking with robust statistics, Pattern Anal. Appl. 2 (1999) 312–320.

[16] G. Schaefer, How useful are colour invariants for image retrieval, in: Proceedings of the Second International Conference, Computer Vision and Graphics, Warsaw, Poland, 2004, pp. 381–386.

[17] G. Finlayson, D. Hordley, G. Schaefer, G. Tian, Illuminant and device invariant colour using histogram equalization, Pattern Recogn. 2 (38) (2005) 179–190.

[18] S.D. Hordley, G.D. Finlayson, G. Schaefer et, G.Y. Tian, Illuminant and Device Invariant Colour Using Histogram Equalisation, Elsevier Science, 2002.

[19] G.D. Finlayson, B. Schiele et, J.L. Crowley, Comprehensive colour image normalization, in: Proceedings of the Fifth European Conference on Computer Vision, 1998, pp. 475–490.

[20] T. Gevers, A. Smeulders, Colour based object recognition, Pattern Recogn. 32 (1999) 453–464.

[21] L. Latecki, V. Rajagopal, A. Gross, Image retrieval and reversible illumination normalization, in: Proceedings of the IS&T/SPIE. Internet Imaging VI, San Jose, USA, 2005.

[22] M. Gouiffés, Apports de la Couleur et des Modéles de Réflexion pour l'Extraction et le Suivi de Primitives, Ph.d. thesis, Université de Poitiers, Poitiers, France, Décembre, 2005.

[23] J. Dargham, Lip detection by the use of neural networks, Artif. Life Robot. 12 (1–2) (2008) 301–306.

[24] T. Gevers, A. Smeulders, Object recognition based on photometric colour invariants, in: Proceedings of the 10th Scandinavian conference on image analysis. Lappeenranta, Finland, 1997, pp. 861–868.

[25] T. Gevers et, H.M.G. Stokman, Classifying of color edges in video into shadow-geometry, highlight, or material transitions, IEEE Trans. Multimed. 5 (2003) 237–243.

[26] T. Gevers, H. Stockman, Robust histogram construction from color invariants for object recognition, IEEE Trans. Pattern Anal. Mach. Intell. 1 (26) (2004) 113–118.

[27] Y. Deng, C. Kenney, M.S. Moore, B.S. Manjunath, Peer group filtering and perceptual color image quantization, IEEE Int. Symp. Circuits Syst. 4 (1999) 21–24.

[28] L.C. Evans, J. Spruck, Motion of level sets by mean curvature: I, J. Differ. Geom. 33 (1991) 635–681.

[29] D. Tschumperlé, Fast anisotropic smoothing of multi-valued images using curvature-preserving PDE's, in: Cahier du GREYC No 05/01, Avril, 2005.

[30] D. Comaniciu, P. Meer, Mean shift: a robust approach toward feature space analysis, IEEE Trans. Pattern Anal. Mach. Intell. 24 (2002) 603–619.

[31] Y. Deng, B.S. Manjunath, Unsupervised segmentation of color-texture regions in images and video, IEEE Trans. Pattern Anal. Mach. Intell. 23 (2000) 800–810.

[32] A. Levinshtein, A. Stere, K.N. Kutulakos, D.J. Fleet, S.J. Dickinson, K. Siddiqi, Turbo pixels: fast super pixels using geometric flows, IEEE Trans. Pattern Anal. Mach. Intell. 31 (2009) 2290–2297.

[33] R. Nock, F. Nielsen, Statistical region merging, IEEE Trans. Pattern Anal. Mach. Intell. 26 (2004) 1452–1458.

[34] B. Sirmacek, C. Unsalan, Building detection from aerial imagery using invariant color features and shadow information, in: Proceedings of the 23rd International Symposium on Computer and Information Sciences, Istanbul, Turkey, 1–5, 2008, pp. 27–29.

[35] C. Gu, J.J. Lim, P. Arbeláez, J. Malik, Recognition using regions, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 1030–1037.

[36] Z. Chao, E. Charles, C. Liming, Image region description using orthogonal combination of local binary patterns enhanced with color information, Pattern Recogn. 46 (2013) 1949–1963.

[37] A. Shahbahrami, D.B.B. Juurlink, Comparison between color and texture features for image retrieval, in: Proceedings of the 19th annual Workshop on Circuitsn Systéms and Signal Processing Veldhoven , Netherlands, 2008.

[38] F. Dornaika, A. Moujahid, Y. El merabet, Y. Ruichek, Building detection from orthophotos using a machine learning approach: an empirical study on image segmentation and descriptors, Expert Syst. Appl. 58 (2016) 130–142.

[39] K. Meskaldji, S. Boucherkha, S. Chikhi, Color quantization and its impact on color histogram based image retrieval accuracy, in: Proceedings of the First International Conference on Networked Digital Technologies, NDT'09, 2009.

[40] N. Loris, L. Alessandra, B. Sheryl, Survey on LBP based texture descriptors for image classification, Expert Syst. Appl. 39 (2012) 3634–3641.

[41] F. Antonio, X.A. Marcos, B. Francesco, Texture description through histograms of equivalent patterns, J. Math. Imag. Vis. 45 (1) (2013) 76–102.

[42] T. Mäenpää, M. Pietikäinen, Classification with color and texture: jointly or separately? Pattern Recogn. Lett. 37 (8) (2004) 1629–1640.

[43] J.Y. Choi, Y.M. Ro, K.N. Plataniotis, Color local texture features for color face recognition, IEEE Trans. Image Process. 21 (3) (2012) 1366–1380.

[44] T. Ojala, M. Pietikäinen, D. Harwood, A comparative study of texture measures

with classification based on feature distributions, Pattern Recogn. 29 (1) (1996) 51–59.
[45] L. Nanni, A. Lumini, S. Brahnam, Local binary patterns variants as texture descriptors for medical image analysis, Artif. Intell. Med. 49 (2) (2010) 117–125.
[46] X. Tan, B. Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, in: Analysis and Modelling of Faces and Gestures. Lecture Notes in Computer Science, volume 4778, Springer, Berlin, 2007, pp. 168–182.
[47] L. Nanni, S. Brahnam, A. Lumini, A local approach based on a local binary patterns variant texture descriptor for classifying pain states, Expert Syst. Appl. 37 (12) (2010) 7888–7894.
[48] M. Heikkilä, M. Pietikäinen, C. Schmid, Description of interest regions with center-symmetric local binary patterns, in: Proceedings of the 5th Indian Conference on Computer Vision, Graphics and Image Processing, Madurai, India, in: Lecture Notes Comput. Sci., 4338, 2006, pp. 58–69.
[49] Y. Huang, Y. Wang, T. Tan, Combining statistics of geometrical and correlative features for 3D face recognition, in: Proceedings of the 17th Conference on British Machine Vision, 2006, pp. 879–888.
[50] B. Zhang, Y. Gao, S. Zhao, J. Liu, Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor, IEEE Trans. Image Process. 19 (2) (2010) 533–544.
[51] V. Ojansivu, J. Heikkila, Blur insensitive texture classification using local phase quantization, ICISP, in: Proceedings of the Lecture Notes in Computer Science LNCS, 5099, 2008, pp. 236–243.
[52] M. Unser, Sum and difference histograms for texture classification, IEEE Trans. Pattern Anal. Mach. Intell. PAMI 8 (1) (1986) 118–125.
[53] M.J. Swain, D.H. Ballard, Color indexing, Int. J. Comput. Vis. 7 (2002) 11–32.
[54] M. Stricker, M. Orengo, Similarity of color images, SPIE Storage Retr. Image Video Databases 3 (1995) 381–392.
[55] T. Kailath, The divergence and bhattacharyya distance measures in signal selection, IEEE Trans. Commun. Technol. 15 (1967) 52–60.
[56] J. Ninga, L. Zhanga, D.Z.C. Wub, Interactive image segmentation by maximal similarity based region merging, Pattern Recogn. 43 (2) (2010) 445–456.
[57] K. Fukunaga, Introduction to Statistical Pattern Recognition, second ed., Academic Press, 1990.

Youssef El MERABET received the M.Sc. degree from Ibn Tofail University, Morocco, in 2008 and the Ph.D. degree from both University of Technology of Belfort-Montbeliard, France, and Ibn Tofail University, Morocco, in 2013. He is now an Assistant Professor in Computer Science and Technology in Ibn Tofail University. His research interests include computer vision and information visualization.

Yassine RUICHEK received respectively his Ph.D in control and computer engineering from the University of Lille, France, in 1997. He was an assistant researcher and teacher from 1998 to 2001. From 2001 to 2007 he was an associate professor at the University of Technology of Belfort- Montbéliard, France. Since 2007, he is a full professor and conducts research activities in multisources/sensors fusion for environment perception and localization. He participated and participates at several labelled projects. His research interests are computer vision, data fusion and pattern recognition. Since 2012, he is the head of the Systems and Transportation Laboratory (SET) of the Research Institute on Transportation, Energy and Society (IRTES).

Saman Ghaffarian received his BS degree in civil engineering from Azad University of Tabriz, Iran, in 2008, and his MSc degree in geomatics engineering from Hacettepe University, Ankara, Turkey, in 2015. He is currently a Ph.D. candidate at the faculty of Geo-Information Science and Earth Observation (ITC), University of Twente,Â Enschede, The Netherlands. His areas of interest are remote sensing, image processing, disaster risk management, spatial economic modeling, automatic object detection, LiDAR point clouds segmentation, object-based image segmentation and classification.

Zineb Samir received the M.Sc. degree from Ibn Tofail University, Morocco, in 2014. She is now a Ph.D. student in Applied physics and Computer Science in Ibn Tofail University. Her research interests include computer vision and structural, thermal and dielectric nanomaterials based on Carbon Nanotubes.

Tarik BOUJIHA received the Ph.D. from the University of Kenitra in 2011. Since, he is an associate professor in the National School of Applied Sciences, Ibn Tofail University. His research topic concerns dynamic texture analysis.

Rochdi MESSOUSSI, received the Ph.D. degree in solid state physics from the University of Nantes, France, in 1990. He received a master degree on the use of technology for learning from the university of Strasbourg in 2001. He is currently professor at Ibn Tofail University, Kénitra, Morocco, where he has been carrying out research about Human Computer Interaction.

Raja Touahni received the Ph.D. from the University of kenitra in 2001. Since, she is a professor in the Ibn Tofail University. Her research topic concerns data analysis and classification.

Abderrahmane SBIHI is the Director of the National School of Applied Sciences (ENSA) of Tangier of Abdelmalek Essaadi University, Morocco, since 2008. From 2001 to 2008, he was the Director of the Laboratory of Telecommunication Systems and Decision Engineering (LASTID) at the Faculty of Sciences, University Ibn Tofail, Kenitra (FSK), Morocco. He is a founding and the President of the Moroccan Society of Classification (MSC). Having a Ph.D. degree in physics in 1995, he has directed a dozen doctoral theses and has published and communicated several researches in the area of multidimensional data analysis and digital image processing.