# Coregistration of Remote Sensing Image Based on Histogram Kernel Predictability

Hugo Carlos ⬤, Ramon Aranda ⬤, Paola A. Mejia-Zuluaga ⬤, Sandra L. Medina-Fernández,
Francisco J. Hernandez-Lopez ⬤, and Miguel Á. Álvarez-Carmona ⬤

*Abstract*—**Registration of remote sensing images has been approached using different strategies; one of the most popular is based on similarity measures. There are different measures of similarity in the literature: Normalized cross-correlation (NCC), mutual information (MI), etc. Normalized mutual information (NMI) has received the most attention in image processing; among the most important limitations are its high computational cost and lack of robustness to strong radiometric changes. For this reason, in this work, we introduce a coregistration approach based on the histogram kernel predictability (HKP). This formulation reduces numerical errors and requires less computing time in comparison to NMI. To the best of our knowledge, this is the first work for registering any remote sensing images by using HKP. Additionally, we propose to use an algorithm based on meta-heuristics called evolutionary centers algorithm, which allows having fewer iterations to solve the registration problem. In addition, we incorporate a parallelization scheme that permits reducing processing times. The results show that our proposal can solve coregistration problems that the NMI cannot solve while obtaining competitive computational times and registration errors comparable with other existing works in the literature. The HKP approach solves most of all the transformations of a set of simulated registration problems, while the NMI, in some cases, only solves half of the registration problems. Moreover, we compare our approach with feature-based methods in real datasets. This research presents an alternative to remote sensing problems where MI has traditionally been used.**

*Index Terms*—**Image registration, kernel predictability (KP), mutual information (MI), remote sensing, satellite images.**

## I. INTRODUCTION

**T**HE coregister of images captured from different devices has generated research interest in applications such as building extraction [1], image classification [2], 3-D city modeling [3], land cover change detection [4], image fusion [5], and radiometric correction [6], [7], among others.

Hugo Carlos is with the Investigadoras e Investigadores por México program in the Centro de Investigacion en Ciencias de Información Geoespacial, Mérida 97307, Méxcico (e-mail: hcarlos@centrogeo.edu.mx).

Ramon Aranda and Francisco J. Hernandez-Lopez are with the Investigadoras e Investigadores por México program in the Centro de Investigación en Matemáticas, Mérida 97307, México (e-mail: arac@cimat.mx; fcoj23@cimat.mx).

Paola A. Mejia-Zuluaga and Sandra L. Medina-Fernández are with the Centro de Investigacion en Ciencias de Información Geoespacial, Ciudad de Méxcico 14240, México (e-mail: paomejia23@gmail.com; sandylizmf@gmail.com).

Miguel Á. Álvarez-Carmona is with the Investigadoras e Investigadores por México program in the Centro de Investigación en Matemáticas, Monterrey 66629, México (e-mail: malvarezc@conacyt.mx).

In the coregistration problem, the idea is to find a geometric transformation that maximizes the alignment between two or more images. The most common options are to estimate the parameters of a similarity, affine, or homography transformation matrix [8]. Coregistration of remote sensing images is challenging because 1) the images have significant geometric and radiometric differences, 2) the presence of noise, and 3) the large size of the images. All these challenges have been addressed using different strategies that include incorporating structural information (e.g., edges) or the use of multiple descriptors [9]. Other approaches make use of statistical measures such as normalized cross-correlation (NCC), mutual information (MI), or normalized mutual information (NMI) to compute the model parameters [10], [11]. In this sense, approaches based on MI have been employed for the coregistration process in the case of multispectral, hyperspectral, and panchromatic images [12], [13], [14]. Also, MI has been successfully applied to the segmentation tasks [15], [16].

Another line of research that has aroused much interest is the coregistration of multimodal images and light detection and ranging (LiDAR) data. For example, Zhu et al. [17] proposed a methodology for registering aerial images and LiDAR; in this work, the authors make efficient use of the structural information through Gabor filters; in the case of MI and NMI, a variety of works address this problem [18], [19]. Even there exist algorithms that use MI and deep learning to register point clouds [20], or applications of MI for calibrating LiDAR and cameras [21].

Despite MI's enormous success in solving different tasks, we can find in the literature information that refers to its limitations. Among the most important are its high computational cost and lack of robustness to strong radiometric changes, its dependence on the selection of parameters, and the optimization method used [22], [23].

The main contribution of this work is to introduce the use of a metric called kernel predictability (KP) for registering remote sensing images. To the best of our knowledge, KP has not been used to register any remote sensing images. The KP has shown superior performance than MI in the image registration process and is more robust regarding the parameters used in its estimation [24]. In particular, we employed a formulation based on histograms called histogram kernel predictability (HKP). This formulation is more robust to numerical errors and requires less computing time than MI [25]. Besides, to address the optimization problem, we propose to use an algorithm based on meta-heuristics called evolutionary centers algorithm (ECA).

ECA allows having fewer iterations to solve the registration problem efficiently. Also, we propose to use single-instruction multiple data (SIMD) and parallel strategies on a multicore architecture to guarantee competitive computational time to incorporate all the available pixels information in the estimation. To evaluate the performance of HKP in the coregistration of remote sensing images, we created a set of synthetic experiments that allows comparing the performance of our proposal in contrast with the NMI formulation. We also use a set of experiments with real datasets, for which we compare our proposal against feature descriptors algorithms encountered in state-of-the-art.

The outline of this article is as follows. Section II presents the works related to the registration process. Section III presents the mathematical background of the KP and its variant HKP. Section IV describes the metaheuristic method used for optimizing the HKP measure. The experimental results are shown in Sections VI and VII for simulated and real datasets, respectively. Finally, Section VIII concludes the article.

## II. REVISITED LITERATURE

In general, remote sensing image registration methods can be classified into two main groups: feature-based and intensity-based methods [26]. For feature-based methods, the objective is to find the correspondences for each descriptor generated in the input images and use these correspondences to minimize an error function. On the other hand, for the image registration process performed by intensity-based methods, the idea is to optimize a function based on the statistical information obtained from the intensity values of the images.

Feature-based methods depend totally on the quality of extracted features and the post-process performed over them (e.g., feature matching and outlier rejection). One of the most successful strategies to address these problems is incorporating edge information (i.e., structural information). For example, [27] combines nonlinear diffusion and phase congruency structural descriptor for the registration of synthetic aperture radar (SAR) and optical images. Paul and Pati [28] propose a method based on scale-invariant feature transform (SIFT) to extract features from the SAR images by considering three important factors: Stability, distinctiveness, and distribution. Xiang et al. [29] propose using the improved phase congruency (PC) model to identify the best features and automatically register optical and SAR images. Zhou et al. [30] use deep learning, structural information, and multiscale strategies to refine the structural descriptors.

Another successful example of using structural information for coregister images is the work presented in [31]. In this work, the authors use the image's structure information to construct the descriptor that can eliminate modal differences and is more suitable for multimode image registration. The idea is to combine the PC describing structural features with the histogram of gradient directions as a feature descriptor, called the histogram of phase congruency (HOPC). Ye et al. [32] extend the idea of HOPC to create a feature descriptor named local histogram of oriented phase congruency (LHOPC). The idea of the LHOPC

method is to solve the radiation differences between multimodal remote sensing (MMRS) images caused by spectral and time changes.

The use of neural networks and deep learning to solve the coregistration problem has attracted the attention of the remote sensing community. For example, Zhang et al. [33] presented a transfer learning approach and local features estimation to register MMRS images. Liu et al. [34] presented a siamese network to deal with the complexity registration of low-altitude remote-sensing images. Although approaches based on deep learning show improvements, the main disadvantage is the high computational cost and the significant number of examples required to train them.

As we mentioned before, in addition to constructing robust descriptors, some works focus on creating computational strategies that allow establishing correspondences (e.g., rejecting outliers). For example, Du et al. [35] propose an approach based on the block and octave constraint SIFT for high-resolution satellite image matching. Lowe [36] proposes employing a database with descriptors of different objects to create a robust correspondence matching algorithm to identify objects in real-time.

Feature-based methods have demonstrated outstanding performance in several cases and practical applications. However, they fail when the noise, geometric deformations, or complex intensity changes cause the descriptor's quality to decrease. Methods based on statistical measures of the image intensity emerged as an alternative to solve these problems.

As we noted before, the formulations based on MI are the most widespread methodologies to solve the coregistration problem. For example, in [18], the registration of optical imagery with LiDAR data is proposed by implementing an alternative approach using image patches and MI. Saidi et al. [37] propose a method for registration of SAR, and the optical images, extracting, and matching features using the MI approach. In [38], a cross-correlation method to obtain coherent points to register interferometric synthetic aperture radar images is proposed; Zhang et al. [39] propose an adapted weighted MI methodology for MMRS image registration. Wu et al. [40] introduce a multimodal continuous ant colony optimization algorithm for remote sensing image registration. Liu et al. [41] present a robust MI registration method for multispectral images with low-resolution and panchromatic images with a high resolution. In [19], a coarse-to-fine registration method is proposed, which includes a super-resolution approach applied to LiDAR data to generate images with the same resolution as the optical image. Finally, Yan et al. [23] present a search algorithm named transfer optimization to avoid getting trapped into local optima by maximizing the MI similarity cost function.

One disadvantage of the intensity-based compared to the feature-based methods is their computational cost. This class of methodologies requires calculating statistics on all the image pixels (or a significant number of samples) to compute the geometric transformation. Besides, the efficiency of these methods is strongly related to the optimization algorithms employed because coregistering the images requires solving nonlinear optimization problems [23], [26].

Due to the above, we address the coregistration problem using the HKP measure. This measure can be estimated efficiently using parallelization strategies, which reduces computation times. Also, using the HKP, it is possible to construct a locally convex similarity function; this property, together with meta-heuristics, solves the problem efficiently [25].

## III. REGISTRATION BASED ON KERNEL PREDICTABILITY

To understand the concept of KP [24], one could do the following exercise: Given a probability distribution $F$, one generates a sample $x_1$ from $F$; the idea is to "guess" the value of $x_1$ from a new sample $x_2$. To evaluate our prediction, it is possible to create a reward function $K(x_1, x_2)$, which is maximal if the distance between $x_1$ and $x_2$ is zero. Given $K(x_1, x_2)$, it is feasible to calculate the expected reward $E[K(X_1, X_2)]$ for a set of samples, where $x_1 \in X_1$ and $x_2 \in X_2$ are values obtained from $F$. From these notions, the authors formulate the KP of a distribution $F$ as

$$
\begin{aligned}
\mathrm{KP}(F) &= E[K(X_1, X_2)] \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(x_1, x_2) dF(x_1) dF(x_2)
\end{aligned} \tag{1}
$$

where $K$ is a kernel function or decreasing function of the distance between $x_1$ and $x_2$. This function measures the predictability of the random variables distributed according to $F$ and weighted by the kernel function $K$. The KP represents a property of the underlying distribution $F$, such as its entropy or its variance. Note that less uncertainty in $F$ implies that the expected reward $E[K(X_1, X_2)]$ will be higher. Besides, KP is a predictability measure, so it behaves inversely compared to entropy, an uncertainty measure.

### A. Estimation of the Kernel Predictability

Equation (1) is a regular statistical functional of degree two (the number of arguments of $K$), and its computation has different estimators. The estimators are based on a sampling set composed of $n$ independent and identically distributed random variables, $\mathbf{X} = \{X_1, X_2, \ldots, X_n\}$ with $X_i \sim F$. The one suggested in [24] is

$$
\widehat{KP} = \frac{4}{n^2} \sum_{i=1}^{n/2} \sum_{j=n/2+1}^{n} K(X_i, X_j). \tag{2}
$$

The selection of this estimator is because the estimator is unbiased, has the lowest computational cost, and presents variance. This last property is helpful in the optimization process used in [24] because it helps to escape from local minima. A natural choice for $K$ is the Gaussian kernel, which is defined as

$$
K(x_1, x_2) = \frac{1}{(2\pi\sigma^2)^{d/2}} \exp\left(-\frac{\|x_1 - x_2\|^2}{2\sigma^2}\right) \tag{3}
$$

with $d$ the dimension of the distribution and $\sigma$ a free parameter which determines the width of the distribution.

### B. Image Registration With Kernel Predictability

In the registration problem, the KP can be applied using the joint distribution of the intensities of a reference image $I_R$ and a transformed image $I_T$ with the following similarity measure:

$$
\mathrm{SKP}(I_T, I_R) = \frac{\mathrm{KP}[p(I_R, I_T)]}{\mathrm{KP}[p(I_T)] + \mathrm{KP}[p(I_R)]} \tag{4}
$$

where $p(I_R, I_T)$ is the joint distribution, and $p(I_R)$ and $p(I_T)$ are the marginal distributions for the images $I_R$ and $I_T$, respectively. The denominator restricts the solution space to avoid anomalous solutions (e.g., all the pixels assigned to a single point). Note that the normalization is similar to the one proposed for NMI [42].

The image registration by *KP* is done by searching the transformation $T(\mathbf{a})$, with $\mathbf{a} \in \mathbb{R}^D$ which maximizes the SKP value. One approximation to (4) can be written as

$$
\widehat{\mathrm{SKP}}(T(\mathbf{a})) = \frac{\widehat{\mathrm{KP}}_J[T(\mathbf{a})]}{\widehat{\mathrm{KP}}_T[T(\mathbf{a})] + \widehat{\mathrm{KP}}_R} \tag{5}
$$

with

$$
\widehat{\mathrm{KP}}_J[T(\mathbf{a})] = \sum_{i=1}^{n/2} \sum_{j=n/2+1}^{n} K_{\sigma_J}(I_J^i, I_J^j)
$$

$$
\widehat{\mathrm{KP}}_T[T(\mathbf{a})] = \sum_{i=1}^{n/2} \sum_{j=n/2+1}^{n} K_{\sigma_M}(I_T^i, I_T^j)
$$

$$
\widehat{\mathrm{KP}}_R = \sum_{i=1}^{n/2} \sum_{j=n/2+1}^{n} K_{\sigma_M}(I_R^i, I_R^j) \tag{6}
$$

$K_{\sigma_J}$ is the kernel to measure the predictability of the joint distribution, and $K_{\sigma_M}$ for the marginal distributions. Note that the constant coefficient in the estimators can be ignored due to normalization. Thus, if Gaussian kernels are used, then

$$
K_{\sigma_J}(I_J^i, I_J^j) = \exp\left(-\frac{\|I_J^i - I_J^j\|^2}{2\sigma_J^2}\right)
$$

$$
K_{\sigma_M}(I^i, I^j) = \exp\left(-\frac{(I^i - I^j)^2}{2\sigma_M^2}\right). \tag{7}
$$

### C. Histogram Kernel Predictability

A new similarity measure based on a discrete version of KP was introduced in [25]. The authors call the novel form of approximating the KP as HKP. The main idea is to approximate the KP by counting bins of histogram frequencies. For instance, given a histogram of $n$ bins, the authors defined the HKP as

$$
\mathrm{HKP}(X) = \sum_{i=1}^{n} B_X(i) \left(\frac{B_X(i) - 1}{N^2}\right) \tag{8}
$$

where $N = \sum_{i=1}^{n} B_X(i)$ is the normalization constant. For two random variables, given the joint histogram $B_{X,Y}(.,.)$, the HKP is given by

$$
\mathrm{HKP}(X, Y) = \sum_{i=1}^{n} \sum_{j=1}^{n} B_{X,Y}(i, j) \left(\frac{B_{X,Y}(i, j) - 1}{N^2}\right) \tag{9}
$$

using this approximation to KP, one can define the similarity measure as

$$\text{SHKP}(I_T, I_R) = \frac{\text{HKP}(I_T, I_R)}{\text{HKP}(I_T) + \text{HKP}(I_R)}. \qquad (10)$$

Among its main advantages are its numerical stability and ease of computation. For example, if the images have the same size, it is possible to calculate the histograms using parallelization strategies based on graphics processing units (GPUs), SIMD, or multicore [43], [44].

For all the KP estimators presented in this section, the estimation quality enhances as the number of samples increases [see (2) and (9)]. To take advantage of this fact, we used different parallelization schemes to use all the pixels in the overlapping regions of the images; this is important because the optimization algorithms that we employ use function evaluations exclusively, then reducing the computation time becomes crucial. In particular, parallelizing the histogram estimation can be done efficiently. Therefore, in this work, we prefer to use the HKP estimator.

### D. Parametric Registration Using SKP-SHKP

The optimization methods for image registration can be classified into two schemes: Convex and meta-heuristic optimization algorithms. In MI, the stochastic gradient descent (SGD) and Powell's method are some of the most popular choices for convex optimization algorithms. In particular, Gómez-García et al. [24] suggest using the following SGD algorithm for the case parametric image coregistration. For a parametric geometric transform $T(\mathbf{a})$ with parameters vector $\mathbf{a} \in \mathbb{R}^D$, the authors use the next update rule

$$\mathbf{a}_{t+1} = \mathbf{a_t} + \alpha \nabla_\mathbf{a} \widehat{\text{SKP}}[T(\mathbf{a}_t)]. \qquad (11)$$

The main drawback of this proposal is that it is necessary to estimate the gradient $\nabla_\mathbf{a} \widehat{\text{SKP}}[T(\mathbf{a}_t)]$ using finite differences because there is no explicit dependency on $T(\mathbf{a})$ in the $\widehat{\text{SKP}}$ estimation (only uses the intensities of the $I_R$ image). Then, the computation of the gradient is done using centered differences

$$\frac{\partial \widehat{\text{SKP}}}{\partial a_i}[T(\mathbf{a}_t)] = \frac{\widehat{\text{SKP}}[T(\mathbf{a}_t + \epsilon_i \mathbf{e}_i)] - \widehat{\text{SKP}}[T(\mathbf{a}_t - \epsilon_i \mathbf{e}_i)]}{2\epsilon_i}$$

where $\epsilon_i$ is a small real value and $\mathbf{e}_i$ is a vector with a one in the $i$th component and zeros in the rest. Of course, there are problems with the numerical approximation of the gradient. The main one is that the similarity must be evaluated twice for each parameter in the transformation. Because every evaluation determines a different overlapping region between the images, to calculate the gradient accurately, the samples used for estimation must lie in the *intersection of all overlapping regions*. Second, selecting values of $\epsilon_i$ can be complicated because the transformation parameters are on different scales. Furthermore, using small values of $\epsilon_i$ can lead to numerical errors during the estimation. Lastly, using numerical methods to estimate the gradient makes escaping from local minima more complicated. To attenuate the latter, Gómez-García et al. [24] suggest using

a pyramidal registration scheme and taking advantage of the estimator's variance (see Section III-A).

## IV. USE OF METAHEURISTICS FOR MULTIMODAL IMAGE REGISTRATION

As discussed in the previous section, various numerical issues are associated with estimating and using numerical derivatives in SKP and SHKP. A possible solution is using algorithms that find the optima using only function evaluations. Different proposals in the literature work only with evaluations of the objective function; among its advantages are not only avoiding the problems associated with the numerical estimations of the derivatives; besides, these algorithms are designed to escape from local minima. In the group of these proposals, perhaps the most popular is the so-called meta-heuristics [45].

For the coregister of multimodal images, we can find articles that use meta-heuristics to unravel the problem. For example, using meta-heuristics for the registration of magnetic resonance (MR) and computed tomography (CT) images, maximizing the MI has proven to be an efficient tool [46]. Dida et al. [46] presented a comparison between the particle swarm optimization (PSO) and the grey wolf optimizer (GWO) for registration of MR and CT images, maximizing the value of the MI. Their results demonstrate that GWO has a better accuracy and processing time than PSO. In [47], a comparison between the differential evolution (DE), PSO, and evolutionary algorithms (EAs) is performed using several benchmark optimization problems. That work found that generally, the DE method has the best performance.

The Bayesian optimization methods are another class of optimization algorithms that use only function evaluations and have had an enormous boom in recent years [48]. Like the meta-heuristics, this class of algorithms does not assume any functional form and can be used as global optimization algorithms. In the case of SKP and SHKP, these algorithms were used successfully in [25].

One of the contributions of this work is the use of meta-heuristics to optimize SHKP. To our knowledge, it is the first time that meta-heuristics has been used to optimize SHKP. Using this type of algorithm avoids the problems associated with the numerical estimation of the gradient and permits escape from local minima. The idea is to use a meta-heuristic that requires fewer function evaluations to converge, with fewer hyperparameters and straightforward implementation. The *Evolutionary Optimization Method Based on Center of Mass* (ECA) algorithm fits the requirements. The following section describes the ECA algorithm, although a complete explanation could be found in [49].

### A. Evolutionary Centers Algorithm (ECA)

ECA belongs to the set of algorithms called evolutionary type. The EAs have successfully solved complex bound-constrained optimization problems [50]. However, the most popular EAs are usually those whose design keeps simple, and their number of parameters is low. Motivated by the above, the authors propose a physics-inspired algorithm based on the center of mass concept
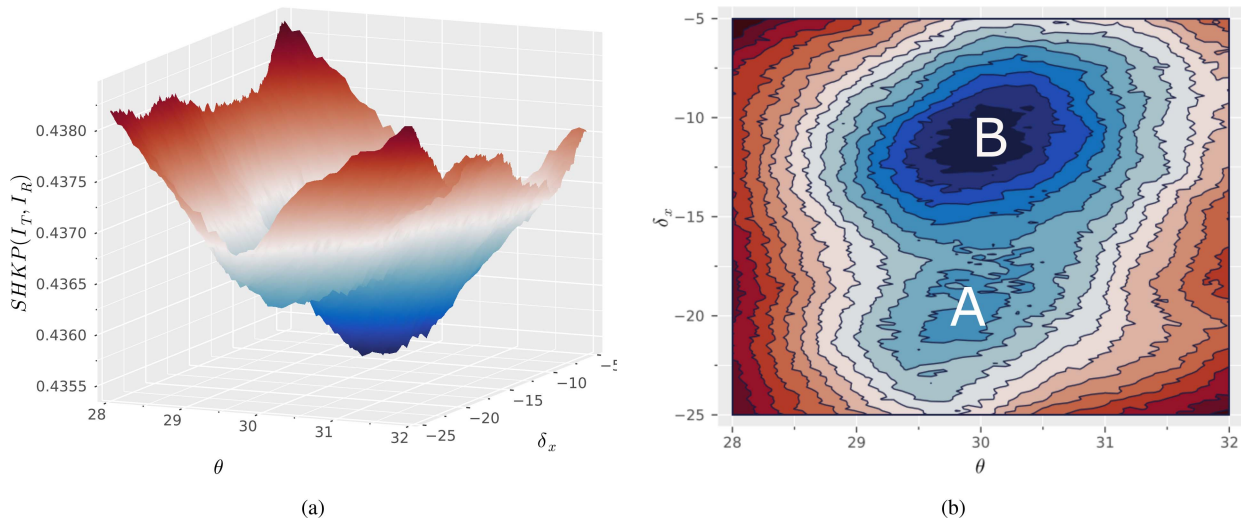
Fig. 1. Surface and contours plots for SHKP function with parameters $[\theta, \delta_x]$. The plots show the existence of numerous local minima around the optimal value.

on a $\mathbb{R}^D$ space for real-parameter single-objective function optimization. The general idea is to promote the creation of an irregular body using $M$ mass points in the current population and then calculate the center of mass to get a new direction for the next population.

Recalling that the idea of ECA is to optimize a function $f : \mathbb{R}^D \to \mathbb{R}$; to do this, ECA has two fundamental steps: 1) estimating the center of mass and 2) the updating rule for the new solutions. Below we describe the fundamental concepts of each step and give hints about their behavior.

*1) Estimation of the Center of Mass:* According to [49], to estimate the center of mass the procedure is the following: Given a set of $N$ solutions $P = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$, a subset $U \subset P$ of $M$ solutions is selected. Then, from $U$, the center of mass is estimated as

$$\mathbf{c} = \frac{1}{W} \sum_{\mathbf{u} \in U} f(\mathbf{u}) \cdot \mathbf{u}, \quad W = \sum_{\mathbf{u} \in U} f(\mathbf{u}). \quad (12)$$

According to the authors, the motivation for using the center of a mass concept consists of translating the population to places where the mass of the entire population is maximum. Theoretically, the set of best solutions should be concentrated near the optimum of the objective function.

*2) Updating Rule for the New Solutions:* To estimate the set of new solutions, the authors proposed the following strategy: First, a random solution $\mathbf{u}_r \in U$ is selected. Then, the update rule for the new solution $\mathbf{h}_i$ is given by

$$\mathbf{h}_i = \mathbf{x}_i + \eta(\mathbf{c} - \mathbf{u}_r). \quad (13)$$

The intuitive idea behind this update rule is the following. Suppose the solutions in the set $P$ are concentrated around an optimum. In that case, the term $(\mathbf{c} - \mathbf{u}_r)$ will be small, which will generate slight perturbations to the current solutions. Eventually, the term will vanish because the set of solutions will be concentrated around the optimum. On the other hand, if the difference $(\mathbf{c} - \mathbf{u}_r)$ is significant, the update rule will allow the exploration of new regions of the objective function, which could

---

**Algorithm 1:** Function ECA.

**Require:** $M, \eta_{\max}$.
  $N \leftarrow 2K * D$.
  Generate and evaluate start population $P$ with $N$ elements.
  **while** the stop criterion is not achieved **do**
    $A = \emptyset$
    **for** $\mathbf{x}_i \in P$ **do**
      Generate $U \subset P$ such that $|U| = M$.
      Estimate $\mathbf{c}$ using (12).
      $\eta \leftarrow rand(0, \eta_{\max})$.
      Select a random solution $\mathbf{u} \in U$.
      $\mathbf{h} \leftarrow \mathbf{x} + \eta(\mathbf{c} - \mathbf{u})$.
      **if** $f(\mathbf{x}) > f(\mathbf{h})$ **then**          $\triangleright \mathbf{h}$ is the best solution
        $A \cup \mathbf{h}$.
      **end if**
    **end for**
    $P \leftarrow$ best elements in $P \cup A$.
  **end while**
  return best solution in $P$.

---

eventually improve the set of current solutions. The $\eta$ factor ponderates the contribution of $(\mathbf{c} - \mathbf{u}_r)$. The random selection of $\eta$ adds an aleatory component to explore the objective function more efficiently.

Algorithm 1 summarizes the essential steps for the implementation of ECAs. Note that a crucial point of the algorithm is the set of initial solutions P. This set is generated over a bounded domain, which limits the search space for solutions. Only local optima will probably be found if the bounden space is small. On the other hand, if the search space is vast, the number of iterations required to find the optimum will be considerable. Regarding the stopping criterion, metaheuristics usually stop after a given number of function evaluations or upon achieving the maximum number of iterations due to the computational cost generated

when evaluating the function. However, it is difficult to estimate a priori the number of iterations needed to converge. In general, these parameters are usually adjusted to achieve a delimited calculation time depending on the nature of the problem.

In Fig. 1(a), we can see the surface of the objective function for SHKP for a pair of hyperspectral images used in our experiments (see Section VI). The surface of the optimization function is convex around the optimum. However, far from the optimum, the function has many local minima, making it challenging to optimize using derivative-based methods. On the other side, meta-heuristics are an excellent choice for their optimization. See the example shown in Fig. 1(b), the simple updating rules proposed in ECA permit escaping from the existing local optima in zone $A$ to the global minima present in $B$.

## V. COREGISTRATION OF REMOTE SENSING IMAGES USING SHKP

Different geometric transformations have been used for the coregistration of remote sensing images. Although the general formulation of SKP and SHKP allows the use of any parametric geometric transformation, we consider that affine transformations are well suited to coregister problems commonly encountered in remote sensing areas. Thus, let $T(\mathbf{a})$ be an *affine transformation* with parameters $\mathbf{a} = [\theta, \lambda_x, \lambda_y, s_x, s_y, \delta_x, \delta_y]$. Then, for a given pixel $\mathbf{p} = [x, y]$, we define the new pixel value $\mathbf{p}_{\text{new}}$ from $T(\mathbf{a})$ as

$$\mathbf{p}_{\text{new}} = (RSH)\mathbf{p} + \mathbf{t} \tag{14}$$

where

$$R = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

$$S = \begin{bmatrix} \lambda_x & 0 \\ 0 & \lambda_y \end{bmatrix}$$

$$H = \begin{bmatrix} 1 & s_x \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ s_y & 1 \end{bmatrix}$$

$$\mathbf{t} = \begin{bmatrix} \delta_x \\ \delta_y \end{bmatrix}$$

where $R$ corresponds to the rotation matrix; $S$ is known as the scaling matrix, $H$ is the shearing matrix, and $\mathbf{t}$ corresponds to the translation vector.

In practice, interpolation algorithms are used when applying geometric transformations to images. Even some works in the literature compare the quality of coregistration based on the interpolation algorithm used [51]. Our experimental phase showed no significant differences between the different interpolation algorithms in the case of SHKP. In this work, a bilinear interpolation was used for all the experiments carried out.
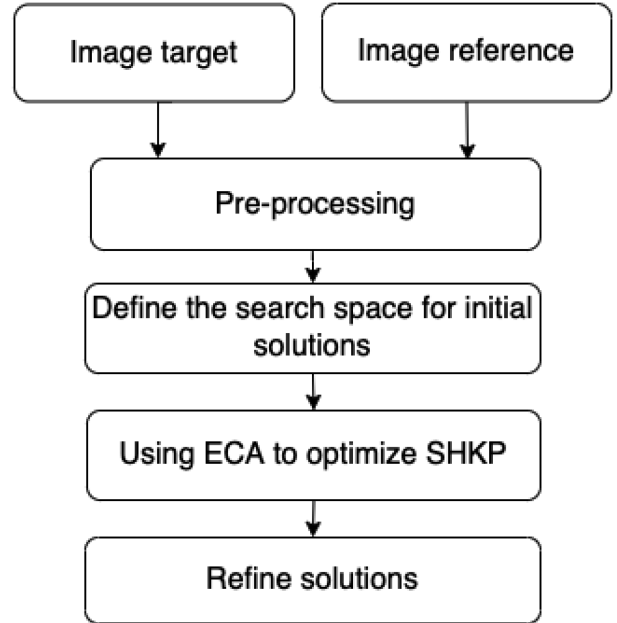


Fig. 2. Flowchart of the proposed image registration methodology.

### A. Proposed Methodology

In Sections III–IV, we present the main concepts of the SHKP and the optimization algorithms that we used in the coregistration processes. This section aims to describe the computational aspects of each of the stages and the details of their implementation. Fig. 2 shows the flowchart of the proposed methodology. Additionally, the source code with all the process can be found here.[1] Below we describe each one of the proposed stages.

*1) Preprocessing:* First, the images are converted to gray-scale and scaled to $[0, 1]$. Then, we use Gaussian filters for noise removal [52]. Although our methodology shows to be robust to the existence of noise, we consider that the use of filters can improve the result.

*2) Set of Initial Solutions:* The ECA algorithm takes as its starting point a set $P$ that contains the initial solutions of the problem (see Algorithm 1). These solutions are usually chosen randomly (over a bounded domain) to explore the cost function adequately. However, for the coregistration problem, several alternatives can be used. For example, for small transformations, it is convenient to create a set of initial solutions close to the identity transformation to reduce the computational cost of the entire optimization process. On the contrary, if they are complex transformations and there is no *a priori* information about the magnitude of the parameters, the solutions can be generated in a broad search space (e.g., $\delta_x$ can be initialized to values within $[-\frac{\text{width}}{2}, \frac{\text{width}}{2}]$). We use this strategy in the registration problems with synthetic data presented in Section VI.

*3) Using ECA to Optimize SHKP:* In Section IV, we describe the theoretical concepts of the ECA algorithm for function optimization. Once the initial search space is defined, the algorithm has two hyperparameters: The number of solutions $M$ used to

---

[1][Online]. Available: https://github.com/hucarlos08/Co-Register-HKP-RS

estimate the center of mass and $\eta_{\max}$ which serves as an upper bound for $\eta$. In the experiments carried out in this work, we set the values of $M = 7$ and $\eta_{\max} = 2$ since they are the values suggested by the authors. However, if the initial search space is wide, increasing the number of samples $M$ used to estimate the center of mass is convenient.

*4) Solution Refinement:* Although the ECA algorithm was designed to find the optimum of the function over a bounded domain, given its stochastic nature, it is still likely to get trapped into a local minimum. To mitigate this effect is possible to attempt to refine the initial solution using the following methodology. Given the *initial solution* **a**, found by ECA for a fixed number of iterations, we can initialize a new instance of the algorithm that uniformly samples points around **a**. If there are *better solutions* in **a**'s neighborhood, it may be possible to explore them. Of course, the number of maximum iterations is adjusted to a number much lower than the initial instance so that the added computational cost is not excessive. In this work, we use this strategy to improve the solutions to the problems presented in Section VII.

### B. Histogram Estimation

As we describe in Section III-C, the estimation of SHKP requires the calculation of the histogram. For example, in the proposal presented in [25], the histogram was estimated from a small set of samples. One of this work's contributions consists in estimating the SHKP with all the available pixels without yielding excessive computational time. In this logic, we consider that the information included in a pixel is valid if it is in the image domain after applying an affine transformation.

In computational terms, selecting which pixels are in the overlap area of images $I_R$ and $I_T$ adds computational cost (and makes it challenging to use SIMD strategies). To avoid these problems, the approach used in this work consists in assigning the pixels outside the image domain to a histogram bin that should not be included in the estimation of SHKP. In our implementation, we generate a histogram of $n + 1$ bins, where the last bin of the histogram corresponds to those pixels outside the image domain. Then, we use the first $n$ bins to estimate SHKP. Although this strategy is straightforward, it permits simultaneous use of multiprocessing and SIMD strategies. Of course, the solution has, as a consequence, an increase in the memory cost, but since the size of the histogram is usually small, this cost is negligible.

### VI. Experimental Results on Simulated Datasets

To evaluate the performance of our proposal, we present the results for a set of synthetic coregistration problems for hyper and multispectral satellite images. For this purpose, we generated a set of 50 random affine transformations $S = \{T(\mathbf{a}_1), T(\mathbf{a}_2), \ldots, T(\mathbf{a}_{50})\}$. The parameters for each affine transformation were uniformly sampled according to the next intervals: $\theta \in [-100, 100]$; $\lambda_x, \lambda_y \in [0.5, 1.5]$, $s_x, s_y \in [0.7, 1.3]$, and $\delta_x, \delta_y \in [-200, 200]$.

The HKP has one hyperparameter: The number of bins $n$. However, as mentioned in [25], the HKP works better with reduced bins like the NMI. For all sets of experiments, we used

$n = 16$ because it was the configuration that generally presented the best results, either HKP or NMI. We analyze the quality of the coregistration in terms of root mean-square-error (RMSE) and mean absolute error (MAE) of *successful registrations*. In this work, we consider a successful registration if the RMSE between the applied and recovery transformations is lower than one-pixel [24]. Additionally, to assess the performance of the optimization algorithm, we analyze the number of iterations to converge and computational time. The experiments were executed on a computer with Intel(R) i7 CPU 2.6 GHz, Ubuntu 18.04 (64-bits), six hyperthreading cores, and 16 GB RAM. All modules were programmed in Julia 1.6 using its capabilities to generate SIMD instructions and multicore parallel schemes [53].

### A. Synthetic Dataset From Hyperspectral Satellite Images

The first set of experiments simulates the coregistration problem of hyperspectral satellite images. The images have a size of $512 \times 512$ pixels and correspond to Iowa, Lake, Santa Cruz, and Vegas.[2] The methodology to create the synthetic dataset was as follows: We obtained the images corresponding to bands one and five for each hyperspectral image. Then, we use the image corresponding to band one as the source image. We apply an affine transformation of the generated set $S$ to this image, and we use the transformed image as the target image. We use the image of band five as the reference image. Fig. 3 illustrates the images generated to simulate the coregistration problem. This set of experiments aims to compare HKP against NMI by using ECA and DE optimizers.

### B. Synthetic Dataset From Multispectral Satellite Images

For the synthetic dataset created from the multispectral images, we used as source images a set of five images of the city of Toulouse, France.[3] The description of the images is shown in Table I. The images were converted to gray scale and scaled to a quarter of their original size and used as the *source* image.

We construct the target and reference images, using a methodology similar to the one presented in [31]. First, we constructed the reference image by applying the next nonlinear tone transfer function to the source image.

$$I(\mathbf{p}) = \exp\left(1 - I(\mathbf{p})\right)^{\gamma} \qquad (15)$$

where $I(\mathbf{p})$ corresponds to the intensity image in the scale [0,1] and $\gamma$ is the scaling factor. In this work, we set $\gamma = 1.35$.

The target image was created by applying an affine transformation and a spatially -varying intensity warped model to the source image [see (16)]. The main idea in (16) is a locally varying intensity field with a mixture of $G$ randomly centered Gaussians to generate the synthetic image. Given a pixel $\mathbf{p}$, the new intensity is calculated as

$$I(\mathbf{p}) = I(\mathbf{p})\left(\alpha + \frac{1}{G}\sum_{i=1}^{G}\exp(-\lambda||\mathbf{p} - \boldsymbol{\mu}_g||^2)\right) \qquad (16)$$

---

[2][Online]. Available: https://serc.carleton.edu/eyesinthesky2/week11/get_to_know_multispectral_imaging.html.

[3][Online]. Available: https://www.intelligence-airbusds.com/imagery/sample-imagery/.
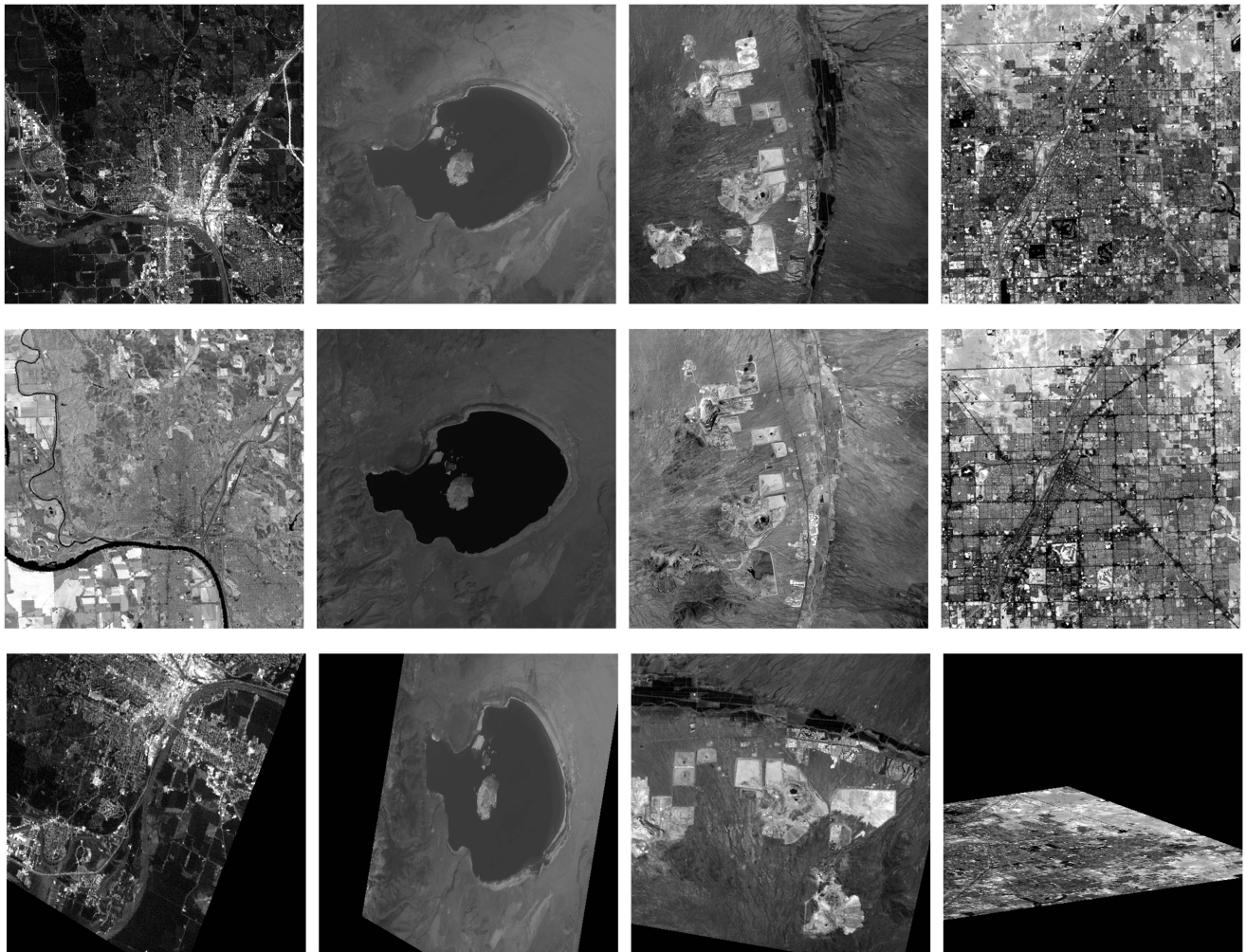
Fig. 3. Each column represents the images used to generate the synthetic dataset. From left to right: Iowa, Lake, Santa Cruz, and Vegas. From top to bottom: source image, reference image, and an example of the target image.

TABLE I
CHARACTERISTICS OF AIRBUS DATASETS FROM TOULOUSE, FRANCE

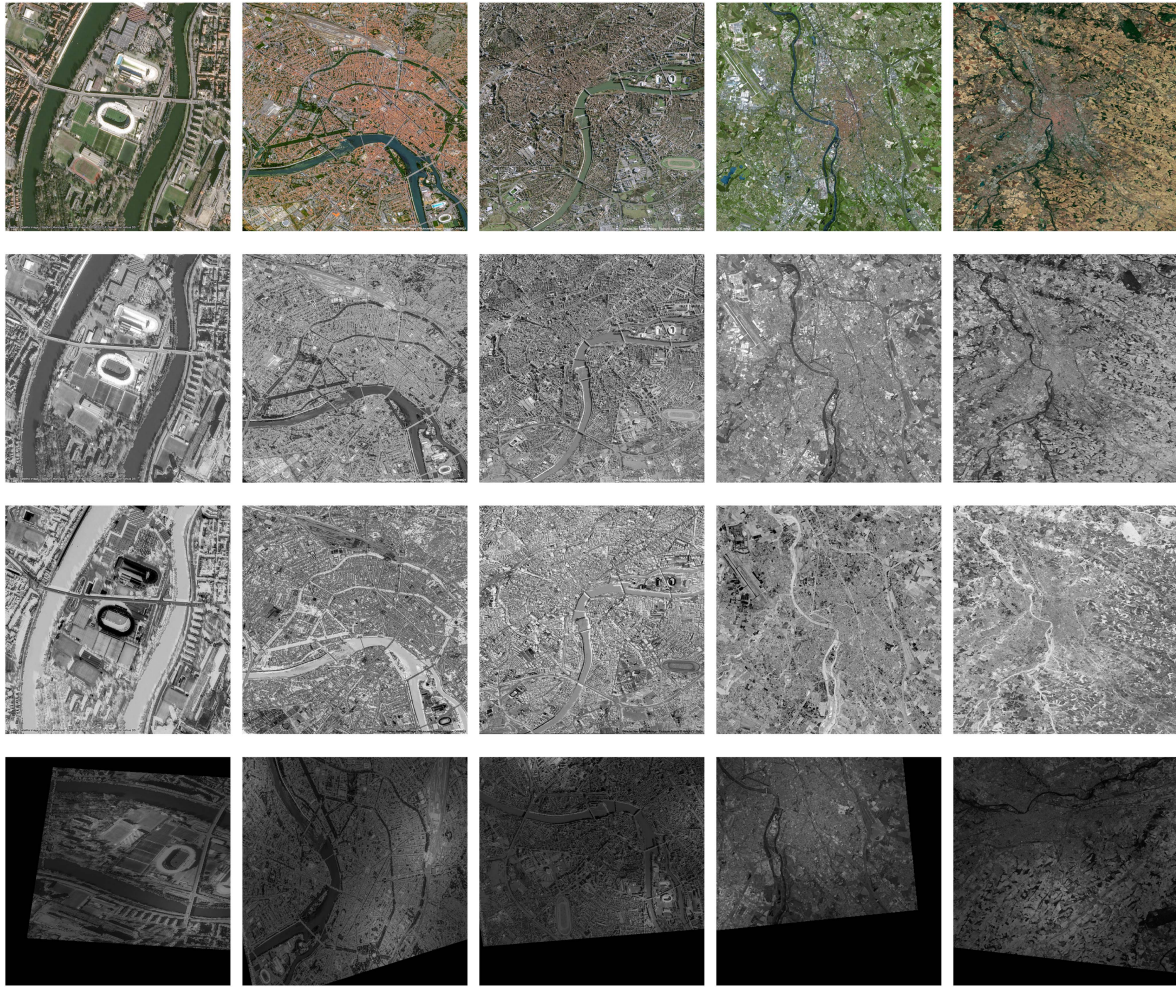| Descriptions of data sets used in the matching experiments | | | | | |
|---|---|---|---|---|---|
| Category | No. | Image pair | Size and GSD | Date | Image characteristic |
| Visible-visible | 1 | Pléiades Neo 1 | 2847 x 2500, 0.30 m | 2022 | Urban area divided by the Garonne river |
| | 2 | Pléiades Neo 2 | 2500 x 2956, 0.30 m | 2021 | Urban area, this image reflects a greater intensity in the colors which accentuates its characteristics |
| | 3 | Pléiades satellite image | 2500 x 2500, 0.50 m | 2016 | Urban area, in the center the municipal stadium of Toulouse next to the Garonne river |
| | 4 | SPOT 5 | 2000 x 2000, 2.5 m | No data | Urban area in Toulouse France |
| | 5 | SPOT 6 | 2500 x 3262, 1.5 m | 2012 | Urban and suburban area, includes the capital of the Occitania region in the south of France and the French department located in the south of the country Haute-Garonne |

Fig. 4. Each column represents the images used to generate the synthetic dataset. From left to right: Pléiades Neo 1, Pléiades Neo 2, Pléiades Satellite Image, SPOT 5, and SPOT 6. From top to bottom: original image, source image, reference image, and an example of the target image.

TABLE II
PROCESSING TIME (MS) AND SPEEDUP FOR HISTOGRAM ESTIMATIONS USING DIFFERENT STRATEGIES AND DIFFERENT NUMBERS OF BINS

| Image size | 16-bins | | | 32-bins | | | 64-bins | | |
|---|---|---|---|---|---|---|---|---|---|
| | Serial (time) | SIMD (time — speedup) | Parallel+SIMD (time — speedup) | Serial (time) | SIMD (time — speedup) | Parallel+SIMD (time — speedup) | Serial (time) | SIMD (time — speedup) | Parallel+SIMD (time — speedup) |
| 256 | 0.481 | 0.396 — 1.214 | 0.257 — 1.871 | 0.491 | 0.406 — 1.209 | 0.271 — 1.811 | 0.493 | 0.417 — 1.182 | 0.165 — 2.987 |
| 512 | 2.023 | 1.795 — 1.127 | 0.914 — 2.213 | 2.044 | 1.651 — 1.238 | 0.599 — 3.412 | 2.110 | 1.695 — 1.244 | 0.540 — 3.907 |
| 1024 | 8.195 | 6.760 — 1.212 | 3.617 — 2.265 | 8.271 | 6.765 — 1.222 | 3.308 — 2.500 | 8.313 | 6.997 — 1.188 | 1.842 — 4.513 |
| 2048 | 32.91 | 29.04 — 1.133 | 14.10 — 2.332 | 32.88 | 27.22 — 1.207 | 11.53 — 2.850 | 33.66 | 28.15 — 1.195 | 8.316 — 4.047 |

where $\alpha$ controls the intensity changes and, $G, \lambda$ controls the intensity spatially varying changes. In this work, we set $G = 3$, $\alpha = 0.3$ and $\lambda = \frac{1}{70^2}$. In Fig. 4, it is possible to observe the source and reference images, and examples of the target image obtained after applying an affine transformation.

### C. Results

This section presents the results of the simulated data. Here, we show the advantages of using parallel strategies in histogram computation and the statistical differences between using HKP and NMI.

*1) Computational Times:* One of the main bottlenecks in evaluating the HKP and NMI is estimating the histogram. To

compare the advantages of using the parallelization scheme, we present the processing time of three different strategies. The first corresponds to an efficient implementation of the histogram calculation without considering any additional parallelization scheme. The second corresponds to an efficient implementation incorporating SIMD instructions. Finally, we present the processing time using the parallelization scheme with multiple cores and SIMD.

Table II presents the processing time average for 1000 runs of the histogram estimation and the speedup, considering different image sizes and number of bins. The speedup is computed by dividing the sequential or serial processing time by the parallel processing time. The results show the advantage of employing parallelization strategies as the image size increases. Our parallel

TABLE III
COMPARISON BETWEEN HKP AND NMI USING DE AND ECA OPTIMIZERS ON
HYPERSPECTRAL IMAGE DATASETS

| Name | Sim | Method | RMSE | MAE | Time | Succ. |
|---|---|---|---|---|---|---|
| Iowa | HKP | DE | 0.182 | 0.161 | 23.126 | 49 |
| | | ECA | 0.143 | 0.129 | 5.916 | 44 |
| | MI | DE | 0.149 | 0.133 | 24.162 | 11 |
| | | ECA | 0.037 | 0.035 | 8.329 | 15 |
| Lake | HKP | DE | 0.167 | 0.147 | 18.256 | 49 |
| | | ECA | 0.200 | 0.176 | 4.124 | 49 |
| | MI | DE | 0.097 | 0.086 | 22.764 | 44 |
| | | ECA | 0.049 | 0.044 | 5.743 | 44 |
| Santa Cruz | HKP | DE | 0.114 | 0.102 | 21.926 | 50 |
| | | ECA | 0.067 | 0.060 | 4.344 | 48 |
| | MI | DE | 0.164 | 0.148 | 22.205 | 12 |
| | | ECA | 0.032 | 0.029 | 6.039 | 12 |
| Vegas | HKP | DE | 0.106 | 0.095 | 24.603 | 50 |
| | | ECA | 0.062 | 0.055 | 5.438 | 46 |
| | MI | DE | 0.207 | 0.180 | 24.517 | 32 |
| | | ECA | 0.011 | 0.010 | 7.053 | 33 |

The estimated median of RMSE (RMSE), MAE (MAE), computational time (TIME), the number of iterations (ITERS), and the number of successful registrations (SUCC).

implementation reaches a speedup up to $4.5\times$ using a multicore CPU and SIMD scheme, launching four threads for an image size of $1024 \times 1024$ pixels and a histogram of 64 bins. Although nowadays, parallel strategies can be implemented quickly, approaches that use multicore strategies require more effort to take care of some details that can generate additional bottlenecks. The advantage of using SIMD algorithms is that they can be easily implemented in many of today's programming languages through macros.

*2) Hyperspectral Datasets:* Table III shows the statistical results for the synthetic datasets created from the hyperspectral images. For these datasets, we not only compared the performance of NMI vs. SHKP but also analyzed the performance of ECA vs. DE.

Table III shows that the combination of NMI and ECA obtains the lowest median for RMSE and MAE. On the other hand, the highest number of successful registrations is obtained by SHKP using either DE or ECA. Furthermore, the optimization algorithm that obtains the shortest computational time is ECA for both SHKP and NMI.

Specifically, Table III shows that NMI solves less than half of the registration problems for the Iowa and Santa Cruz images. For the Vegas dataset, SHKP is superior at around 36% in the successful registration score. For the Lake image, there is no practical difference between NMI and SHKP. In combination with SHKP, DE has a slightly higher number of successful registrations than that obtained by ECA with SHKP. The disadvantage is that it is computationally expensive, with a computation time $4\times$ greater than ECA with SHKP.

*3) Multispectral Datasets:* Table IV shows the results for the datasets created from the Toulouse images. For this set of experiments, we only evaluated the behavior of SHKP and NMI using ECA as an optimization method; this is because previous experiments showed the additional computational cost of using DE.

In the case of Pleiades Neo 1 images, there is a slight superiority in the number of successful registrations obtained by SHKP and maintaining a lower computation time. However, the RMSE and MAE metrics are lower in the case of NMI. For the Pléiades Neo 2 image, both SHKP and NMI obtained some successful

registrations slightly above 50%, this dataset being the one with the worst results.

NMI gets insufficient successful registrations for Pléiades Satellite Image, SPOT 5, y SPOT 6 (close to 10%). The images are highly textured, and NMI has problems coregistering images with these characteristics. Furthermore, the computational time obtained by NMI is almost double that obtained by SHKP. That is because the number of iterations and evaluations of the function is lower in the case of SHKP. The number of successful registrations obtained by SHKP is close to 90%, which shows its superiority in this class of coregistration problems.

It is important to note that the number of successful registrations obtained is higher than those reported in [24], which shows the advantages of using SHKP with ECA for the coregistration problem. The average time to complete the registration is around 20 s. The computational time is significantly less than that reported in [25]; this improvement is due to the use SIMD and multicore CPU parallel scheme to calculate the histograms. The obtained number of successful registrations, the RMSE, and MAE show that our proposal has a competitive performance compared to NMI, and other existing works in the literature [54], [55].

## VII. RESULTS AND DISCUSSION ON REAL DATASET

We evaluate our proposal using different kinds of MMRS images. These datasets correspond to common problems found in coregistration problems. A complete description of the employed images is presented in Table V. We compare our proposal with the algorithms: HOPC (https://github.com/yeyuanxin110/HOPCHOPC), Channel Features of Orientated Gradients (https://github.com/yeyuanxin110/CFOGCFOG) [56], and Co-occurrence Filter Space Matching (https://skyearth.org/publication/project/CoFSM/CoFSM) [57]. All implementations used in this comparison are public and can be found at the links provided by the authors.

The procedure to find the optimal transformation for descriptor-based methodologies (CFOG, CoFSM, and HOPC) was the following. First, we use the Harris corner detector to determine an initial set of control points (CP). Subsequently, we used different methodologies to generate the descriptors and found the correspondences. We use a double-checking strategy to determine the correspondences: We found the descriptors $(i, j)$ such that $i$th descriptor in image $I_T$ has $j$th descriptor in image $I_R$ as the best match and vice versa. Finally, we compute an optimal affine transformation between two 2-D point sets using an algorithm based on the random samples consensus and using a nonlinear optimizer [58].

Table VI shows the quantitative results for each methodology. The results show that our proposal has a performance comparable to other state-of-the-art algorithms. For example, we are obtaining the best results for Lidar-Optical images. In the rest of the images, our proposal has competitive performance. Our worst result corresponds to the Optical-SAR images, where our methodology cannot solve the problem satisfactorily due to the existence of local textures. Our proposal does not use local information; on the contrary, we use the total pixels of the image to solve the registration problem; this is a disadvantage

TABLE IV
COMPARISON BETWEEN HKP AND NMI USING ECA OPTIMIZER ON MULTISPECTRAL IMAGE DATASETS

| Name | Sim | RMSE | MAE | Time(s) | Iterations | Function calls | Succ. |
|---|---|---|---|---|---|---|---|
| Pléiades Neo 1 | HKP | 0.018 | 0.016 | 9.714 | 79 | 2762 | 46 |
| | MI | 0.002 | 0.002 | 12.919 | 100 | 3479 | 38 |
| Pléiades Neo 2 | HKP | 0.007 | 0.007 | 13.262 | 86 | 3003 | 29 |
| | MI | 0.001 | 0.001 | 15.276 | 107 | 3741 | 22 |
| Pléiades satellite image | HKP | 0.005 | 0.004 | 26.278 | 96 | 3360 | 40 |
| | MI | 0.001 | 0.001 | 55.478 | 226 | 7900 | 3 |
| SPOT 5 | HKP | 0.019 | 0.017 | 6.962 | 85 | 2947 | 40 |
| | MI | 0.003 | 0.003 | 12.193 | 150 | 5252 | 8 |
| SPOT 6 | HKP | 0.009 | 0.008 | 13.016 | 84 | 2934 | 40 |
| | MI | 0.002 | 0.002 | 26.065 | 167 | 5842 | 2 |

The estimated median of RMSE (RMSE), MAE (MAE), computational time (TIME), number of iterations (ITERS),
function calls, and number of successful registrations (SUCC).

TABLE V
DESCRIPTIONS OF DATASETS USED IN THE MATCHING EXPERIMENTS

| Category | No. | Image pair | Size and GSD | Date | Image characteristic |
|---|---|---|---|---|---|
| Optical-LiDAR | 1 | LiDAR intensity | 600×600, 2m | Oct, 10 | Urban areas with high buildings, significant local distortions, a temporal difference of 12 months, and significant noise in the LiDAR data |
| | | WorldView2 visible | 600×600, 2m | Oct, 11 | |
| | 2 | LiDAR height | 524×524, 2.5m | Jun, 12 | Urban areas, and large difference in intensity characteristics |
| | | Airborne visible | 524×524, 2.5m | Jun, 12 | |
| Optical-SAR | 1 | Image from Google Earth | 528×524, 3m | Nov, 07 | Urban area, and significant noise in the SAR data 12/2007 |
| | | TerraSAR-X | 528×524, 3m | Dec, 07 | |
| Visible-Infrared | 1 | Daedalus visible | 512×512, 0.5m | Apr, 00 | Urban area |
| | | Daedalus infrared | 512×512, 0.5m | Apr, 00 | |

TABLE VI
RESULTS FOR REAL DATASETS PRESENTED IN FIG. 5

| Dataset | Method | RMSE | MAE |
|---|---|---|---|
| **Lidar-Optical** | SHKP | *1.71* | 1.44 |
| | CFOG | 1.83 | 1.72 |
| | HOPC | 2.24 | *1.38* |
| | CoFSM | 1.88 | 1.72 |
| **Optical-Lidar1** | SHKP | *1.12* | *1.0* |
| | CFOG | 1.88 | 1.33 |
| | HOPC | 1.98 | 1.82 |
| | CoFSM | 1.88 | 1.72 |
| **Optical-SAR** | SHKP | 3.1 | 2.9 |
| | CFOG | 2.43 | 2.17 |
| | HOPC | 2.98 | 2.92 |
| | CoFSM | *1.96* | *1.82* |
| **Visible-infrared2** | SHKP | 0.6 | 0.51 |
| | CFOG | *0.22* | *0.1* |
| | HOPC | 0.38 | 0.28 |
| | CoFSM | 1.78 | 1.61 |

The table shows the RMSE and MAE for the SHKP, CFOG,
HOPC, and CoFSM methods.
Italics represent the best values for each metric by dataset.

the result of our methodology can be considered adequate (less than 1.0 pixel). However, the HOPC and CFOG methodologies are the ones that have the best performance. The difference in the error obtained when coregistering the images can be explained if we consider that our registration method does not incorporate edge (or structural) information in contrast with the descriptors employed in this comparison. Finally, it is essential to point out that our methodology is aimed at solving the coregistration problem using affine transformations. Although extending our formulation to other types of geometric transformations is possible, it may be necessary to readjust the optimization method or the hyperparameters used. The above is an advantage of feature description-based methods; finding the appropriate geometric transformation is independent of constructing and matching descriptors.

Fig. 5 shows the images with the qualitative results of the coregister obtained by SHKP. In this figure, we can see in each column: 1) the target and reference image with the CP employed; 2) the intensity difference images before and after coregistration; and 3) the chess-board images before and after the coregistration. In general, the intensity difference images show a significant improvement in the details of the coregistered images. In particular, for the images that use LiDAR data and visible images, an improvement in the details of the coregistered images is noted. In the case of the visible–SAR image pair, there is a significant improvement in the alignment of the structures, and the details in the highly textured areas are clearly defined. Finally, in the pair of visible–infrared images, one can see the improved alignment of the edges that define the buildings.

in images with "local" textures because local information is not used efficiently. For this class of images, robust methodologies based on descriptors are a better alternative, as shown by the results obtained by CoFSM. Finally, for visible–infrared images,
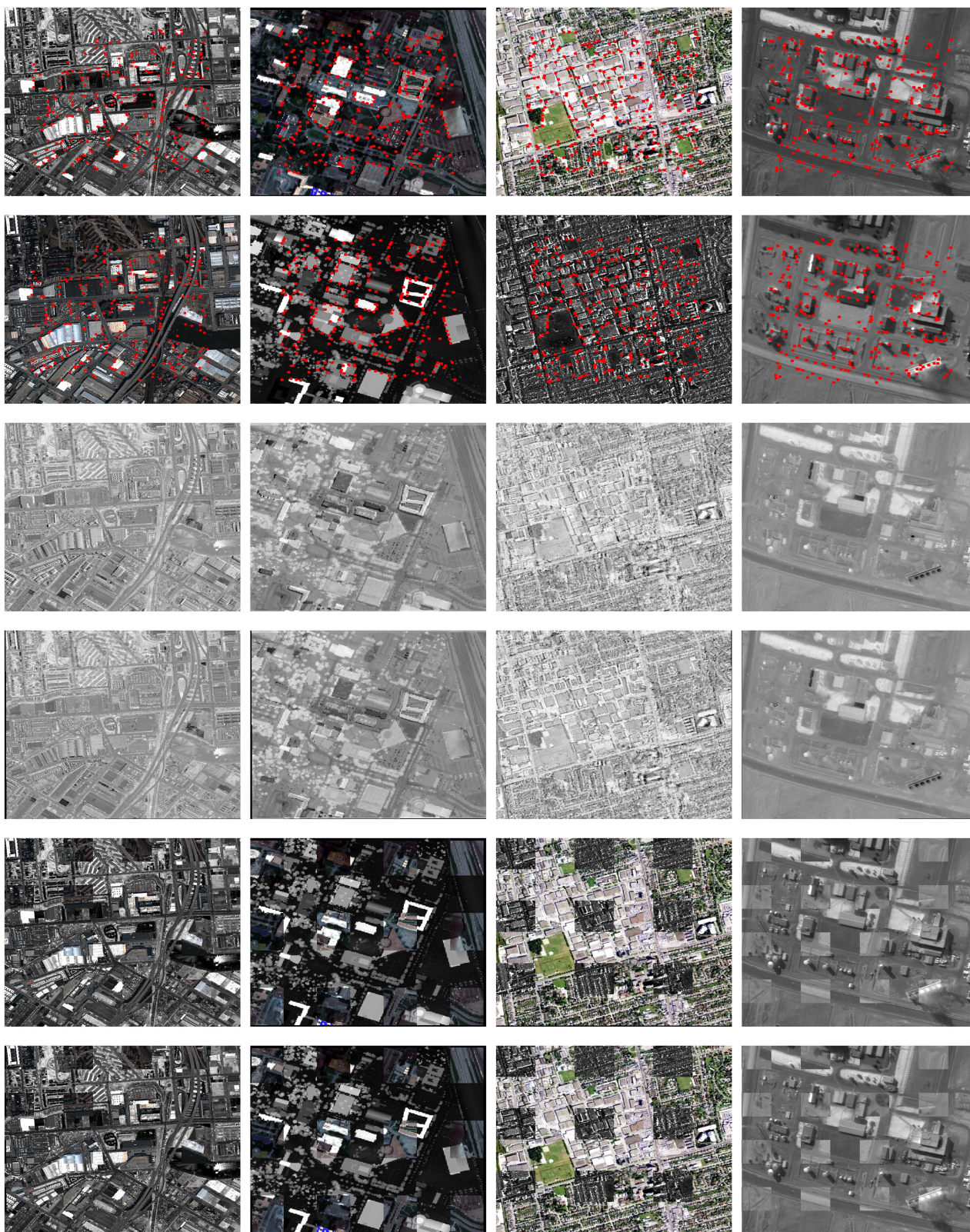
Fig. 5.    Real dataset results. Each column represents a coregistration problem. From left to right: LiDAR-Optical, Optical-LiDAR, Optical-SAR, and visible–infrared. From top to bottom: target image with HOPC CPs, reference image with HOPC CPs, difference image before coregistration, difference image after coregistration, the chess-board image before coregister, and the chess-board image after coregister.

## VIII. Conclusion

The present work shows the successful use of SHKP for remote sensing image coregistration. The results illustrate SHKP's ability to solve registry problems that MI cannot solve. As in other investigations, it is possible to observe that the optimization method strongly impacts the quality of the register. The results show that NMI has a lower RMSE than HKP; however, the number of successful registrations is significantly lower. This effect is seen quite clearly in multispectral image experiments. The suggested optimization algorithm, ECA, benefits both NMI and HKP, having shorter computation times than DE without compromising the RMSE. These results show that HKP could become an alternative in remote sensing problems where MI has traditionally been used.

Additionally, our experiments show that parallelization strategies reduce computation time significantly, allowing processing of all the pixel information available; for example, images with size $512 \times 512$ obtain an average time of around 3 s.

As feature work, it is convenient to consider a fully implemented pipeline in GPUs to reduce computational time and allow the coregister of larger images. It is also essential to incorporate structural information from the images through edges or other features to address the registration problem when the texture in both images is homogeneous.

## References

[1] L. Luo, P. Li, and X. Yan, "Deep learning-based building extraction from remote sensing images: A comprehensive review," *Energies*, vol. 14, no. 23, 2021, Art. no. 7982.

[2] D. Sowmya, P. Deepa Shenoy, and K. Venugopal, "Remote sensing satellite image processing techniques for image classification: A comprehensive survey," *Int. J. Comput. Appl.*, vol. 161, no. 11, pp. 24–37, 2017.

[3] R. Kumar Mishra and Y. Zhang, "A review of optical imagery and airborne Lidar data registration methods," *Open Remote Sens. J.*, vol. 5, no. 1, pp. 54–63, 2012.

[4] L. ZhiYong, T. Liu, J. A. Benediktsson, and N. Falco, "Land cover change detection techniques: Very-high-resolution optical images: A review," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 1, pp. 44–63, Mar. 2022.

[5] H. Ghassemian, "A review of remote sensing image fusion methods," *Inf. Fusion*, vol. 32, pp. 75–89, 2016.

[6] A. Moghimi, T. Celik, A. Mohammadzadeh, and H. Kusetogullari, "Comparison of keypoint detectors and descriptors for relative radiometric normalization of bitemporal remote sensing images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 4063–4073, 2021.

[7] A. Moghimi, A. Sarmadian, A. Mohammadzadeh, T. Celik, M. Amani, and H. Kusetogullari, "Distortion robust relative radiometric normalization of multitemporal and multisensor remote sensing images using image features," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2021, Art. no. 5400820.

[8] R. Szeliski, "Image alignment and stitching: A tutorial," *Found. Trends Comput. Graph. Vis.*, vol. 2, no. 1, pp. 1–104, 2006.

[9] C. Leng, H. Zhang, B. Li, G. Cai, Z. Pei, and L. He, "Local feature descriptor for image matching: A survey," *IEEE Access*, vol. 7, pp. 6424–6434, 2018.

[10] A. Alba, R. M. Aguilar-Ponce, J. F. Vigueras-Gómez, and E. Arce-Santana, "Phase correlation based image alignment with subpixel accuracy," in *Proc. Mex. Int. Conf. Artif. Intell.*, 2012, pp. 171–182.

[11] A. Dame and E. Marchand, "Second-order optimization of mutual information for real-time image registration," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 4190–4203, Sep. 2012.

[12] S. M. Moorthi, D. Dhar, and R. Sivakumar, "Co-registration of LISS-4 multispectral band data using mutual information-based stochastic gradient descent optimization," *Curr. Sci.*, vol. 113, no. 5, pp. 877–888, 2017.

[13] Y. Han, J. Yeom, Y. Kim, and Y. Kim, "Automatic image registration of multi-temporal Kompsat-2 images in agricultural areas," *Res. J. Chem. Environ.*, vol. 18, no. 6, pp. 60–65, 2014.

[14] C. Platias, M. Vakalopoulou, and K. Karantzalos, "Automatic MRF-based registration of high resolution satellite video data," *ISPRS Ann. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, vol. 3, no. 1, pp. 121–128, 2016.

[15] M. A. Hossain, X. Jia, and M. Pickering, "Subspace detection using a mutual information measure for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 2, pp. 424–428, Feb. 2014.

[16] L. Liang, S. Ning, and W. Yan, "Segment-based remote sensing image change detection using normalized mutual information," *Remote Sens. Inf.*, vol. 6, pp. 18–22, 2011.

[17] B. Zhu, Y. Ye, L. Zhou, Z. Li, and G. Yin, "Robust registration of aerial images and LiDAR data using spatial constraints and Gabor structural features," *ISPRS J. Photogrammetry Remote Sens.*, vol. 181, pp. 129–147, 2021.

[18] E. G. Parmehr, C. S. Fraser, C. Zhang, and J. Leach, "Automatic registration of optical imagery with 3D LiDAR data using local combined mutual information," *ISPRS Ann. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, vol. 88, pp. 28–40, 2014.

[19] T. H. Nguyen, S. Daniel, D. Gueriot, C. Sintes, and J.-M. Le Caillec, "Coarse-to-fine registration of airborne LiDAR data and optical imagery on urban scenes," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 3125–3144, 2020.

[20] X. Pan, X. Ji, and S. Cheng, "3D point cloud registration based on cascaded mutual information attention network," in *Proc. 25th Int. Conf. Pattern Recognit.*, 2021, pp. 10644–10649.

[21] P. Jiang, P. R. Osteen, and S. Saripalli, "Calibrating LiDAR and camera using semantic mutual information," *Comput. Res. Repository*, vol. abs/2104.12023, 2021. [Online]. Available: https://arxiv.org/abs/2104.12023

[22] Y. Ye, C. Yang, B. Zhu, L. Zhou, Y. He, and H. Jia, "Improving coregistration for sentinel-1 SAR and sentinel-2 optical images," *Remote Sens.*, vol. 13, no. 5, 2021, Art. no. 928.

[23] X. Yan, Y. Zhang, D. Zhang, N. Hou, and B. Zhang, "Registration of multimodal remote sensing images using transfer optimization," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 12, pp. 2060–2064, Dec. 2020.

[24] H. F. Gómez-García, J. L. Marroquín, and J. Van Horebeek, "Image registration based on kernel-predictability," *Comput. Vis. Image Understanding*, vol. 112, no. 2, pp. 160–172, 2008.

[25] I. Reducindo, E. R. Arce-Santana, D. U. Campos-Delgado, J. F. Vigueras-Gómez, and A. Alba, "An exploration of multimodal similarity metrics for parametric image registration based on particle filtering," in *Proc. 8th Int. Conf. Elect. Eng., Comput. Sci. Automat. Control*, 2011, pp. 1–6.

[26] S. Paul and U. C. Pati, "A comprehensive review on remote sensing image registration," *Int. J. Remote Sens.*, vol. 42, no. 14, pp. 5396–5432, 2021. doi: 10.1080/01431161.2021.190698.

[27] J. Fan, Y. Wu, M. Li, W. Liang, and Y. Cao, "SAR and optical image registration using nonlinear diffusion and phase congruency structural descriptor," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5368–5379, Sep. 2018.

[28] S. Paul and U. C. Pati, "SAR image registration using an improved SAR-sift algorithm and Delaunay-triangulation-based local matching," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 8, pp. 2958–2966, Aug. 2019.

[29] Y. Xiang, R. Tao, F. Wang, H. You, and B. Han, "Automatic registration of optical and SAR images via improved phase congruency model," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 5847–5861, 2020.

[30] L. Zhou, Y. Ye, T. Tang, K. Nan, and Y. Qin, "Robust matching for SAR and optical images using multiscale convolutional gradient features," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2021, Art. no. 4017605.

[31] Y. Ye, J. Shan, L. Bruzzone, and L. Shen, "Robust registration of multimodal remote sensing images based on structural similarity," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2941–2958, May 2017.

[32] Y. Ye, J. Shan, S. Hao, L. Bruzzone, and Y. Qin, "A local phase based invariant feature for remote sensing image matching," *ISPRS J. Photogrammetry Remote Sens.*, vol. 142, pp. 205–221, 2018.

[33] J. Zhang, W. Ma, Y. Wu, and L. Jiao, "Multimodal remote sensing image registration based on image transfer and local features," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 8, pp. 1210–1214, Aug. 2019.

[34] Y. Liu, X. Gong, J. Chen, S. Chen, and Y. Yang, "Rotation-invariant Siamese network for low-altitude remote-sensing image registration," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 5746–5758, 2020.

[35] S. Du, M. Wang, and S. Fang, "Block-and-octave constraint sift with multi-thread processing for VHR satellite image matching," *Remote Sens. Lett.*, vol. 8, no. 12, pp. 1180–1189, 2017, doi: 10.1080/2150704X.2017.1368097.

[36] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[37] F. Saidi, J. Chen, and P. Wang, "A refined automatic co-registration method for high-resolution optical and SAR images by maximizing mutual information," in *Proc. IEEE Int. Conf. Signal Image Process.*, 2016, pp. 231–235.

[38] D. Fang, X. Lv, Y. Yun, and F. Li, "An Insar fine registration algorithm using uniform tie points based on Voronoi diagram," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 8, pp. 1403–1407, Aug. 2017.

[39] J. Zhang, M. Zareapoor, X. He, D. Shen, D. Feng, and J. Yang, "Mutual information based multi-modal remote sensing image registration using adaptive feature weight," *Remote Sens. Lett.*, vol. 9, no. 7, pp. 646–655, 2018, doi: 10.1080/2150704X.2018.1458343.

[40] Y. Wu, W. Ma, Q. Miao, and S. Wang, "Multimodal continuous ant colony optimization for multisensor remote sensing image registration with local search," *Swarm Evol. Comput.*, vol. 47, pp. 89–95, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210650217301864

[41] D. Liu, H. Mansour, and P. T. Boufounos, "Robust mutual information-based multi-image registration," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 915–918.

[42] J. P. Pluim, J. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: A survey," *IEEE Trans. Med. Imag.*, vol. 22, no. 8, pp. 986–1004, Aug. 2003.

[43] A. Shahbahrami, B. Juurlink, and S. Vassiliadis, "SIMD vectorization of histogram functions," in *Proc. IEEE Int. Conf. Appl.-Specific Syst., Architectures Processors*, 2007, pp. 174–179.

[44] T. Henriksen, S. Hellfritzsch, P. Sadayappan, and C. Oancea, "Compiling generalized histograms for GPU," in *Proc. SC20: Int. Conf. High Perform. Comput., Netw., Storage Anal.*, 2020, pp. 1–14.

[45] H. Malik, A. Iqbal, P. Joshi, S. Agrawal, and F. I. Bakhsh, *Metaheuristic and Evolutionary Computation: Algorithms and Applications*. Berlin, Germany: Springer, 2021.

[46] H. Dida, F. Charif, and A. Benchabane, "A comparative study of two meta-heuristic algorithms for MRI and CT images registration," in *Proc. 3rd Int. Conf. Inf. Commun. Technol.*, 2020, pp. 411–415.

[47] J. Vesterstrom and R. Thomsen, "A comparative study of differential evolution, particle swarm optimization, and evolutionary algorithms on numerical benchmark problems," in *Proc. Congr. Evol. Comput.*, 2004, vol. 2, pp. 1980–1987.

[48] P. I. Frazier, "Bayesian optimization," in *Recent Advances in Optimization and Modeling of Contemporary Problems*. INFORMS TutORials in Operations Research, 2018, pp. 255–278.

[49] K. Deep, M. Jain, and S. Salhi, *Decision Science in Action: Theory and Applications of Modern Decision Analytic Optimisation*. Springer, 2019.

[50] A. N. Sloss and S. Gustafson, "2019 evolutionary algorithms review," in *Genetic Programming Theory and Practice XVII*, Berlin, Germany: Springer, 2020, pp. 307–344.

[51] J. P. Pluim, J. A. Maintz, and M. A. Viergever, "Interpolation artefacts in mutual information-based image registration," *Comput. Vis. Image Understanding*, vol. 77, no. 2, pp. 211–232, 2000.

[52] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1989.

[53] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, "Julia: A fresh approach to numerical computing," *SIAM Rev.*, vol. 59, no. 1, pp. 65–98, 2017.

[54] K. Yang, A. Pan, Y. Yang, S. Zhang, S. H. Ong, and H. Tang, "Remote sensing image registration using multiple image features," *Remote Sens.*, vol. 9, no. 6, 2017, Art. no. 581.

[55] L. Rasmy, I. Sebari, and M. Ettarid, "Automatic sub-pixel co-registration of remote sensing images using phase correlation and harris detector," *Remote Sens.*, vol. 13, no. 12, 2021, Art. no. 2314.

[56] Y. Ye, L. Bruzzone, J. Shan, F. Bovolo, and Q. Zhu, "Fast and robust matching for multimodal remote sensing image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 9059–9070, Nov. 2019.

[57] Y. Yao, Y. Zhang, Y. Wan, X. Liu, X. Yan, and J. Li, "Multi-modal remote sensing image matching considering co-occurrence filter," *IEEE Trans. Image Process.*, vol. 31, pp. 2584–2597, 2022.

[58] G. Bradski and A. Kaehler, "Opencv," *Dr Dobb's J. Softw. Tools*, vol. 3, 2000, Art. no. 2.

**Hugo Carlos** received the Ph.D. degree in computer science from Centro de Investigación en Matemáticas (CIMAT), Mérida, Mexico, in 2018.

He is a CONACyT-Researcher with CentroGeo, Ciudad de Méxcico, Mérida, Mexico. His research interests include geointelligence, remote sensing applications, computer vision, optimization, and machine learning.

Dr. Carlos is a member of the Mexican National System of Researchers (SNI).

**Ramon Aranda** received the Ph.D. degree in computer science from the Centro de Investigación en Matemáticas (CIMAT), Mérida, Mexico, in 2016.

He is a CONACyT-Researcher with CIMAT, Mérida, Mexico. His research interests include tourism data analysis, computer vision, optimization, machine learning, and data science.

Dr. Aranda is a member of the Mexican National System of Researchers (SNI).

**Paola A. Mejia-Zuluaga** received the Master's degree in geospatial information sciences, in 2021, from CentroGeo, Ciudad de Méxcico, México, where she is currently working toward the Ph.D. degree in geospatial information sciences.

Her research interests include remote sensing, computer vision, the implementation of evolutionary algorithms, and machine learning focused on tasks of classification, estimation, and study of different phenomena in the territory.

**Sandra L. Medina-Fernández** received the Bachelor's degree in geomatics engineering from the Faculty of Engineering, Universidad Nacional Autonoma de México, Mexico, in 2017, and the Master's degree in geospatial information sciences from the Centro de Investigación en Ciencias de Información Geoespacial (CentroGeo), Ciudad de México, México, in 2021.

Her research interests include topography, bathymetry, remote sensing (PR), and geographic information systems (GIS).

**Francisco J. Hernandez-Lopez** received the M.Sc. and D.Sc. degrees in computer science from the Centro de Investigación en Matemáticas (CIMAT), Mérida, Mexico, in 2009 and 2014, respectively.

Since 2014, he has been with Computer Science Department, CIMAT. His research interests include computer vision, machine learning, and the development of efficient algorithms using parallel computing to process and analyze video sequences.

Hernandez-Lopez is a Fellow of the Mexican National System of Researchers (SNI).

**Miguel Á. Álvarez-Carmona** received the Master's and doctorate degrees in computational sciences from the National Institute of Astrophysics, Optics, and Electronics, Puebla, México, in 2014 and 2019, respectively.

His research interests include the application of artificial intelligence to tourism.

Dr. Álvarez-Carmona is a member of the Mexican National System of Researchers, the Mexican Association of Natural Language Processing, and the Mexican Association of Artificial Intelligence.