# Vision-Based Robot Positioning by an Exact Distance Between Histograms

Francesc Serratosa[1] & Alberto Sanfeliu [2]
[1]Universitat Rovira i Virgili, [2]Universitat Politècnica de Catalunya
francesc.serratosa@urv.cat, sanfeliu@iri.upc.edu

## Abstract

*Most vision-based robot positioning techniques rely on analytical formulations of the relationship between the robot pose and the projected image coordinates of several geometric features of the observed scene. This usually requires that several simple features such as points, lines or circles be visible in the image and be properly extracted.*

*In this paper, we present a method to compare images (scenes that the robot has learned) based on a fast and exact distance between histograms. In contrast to the methods described before, our method is faster and with less storage space do to the images do not need to be segmented and only a lossless description of the histograms are stored in the data base.*

## 1. Introduction

Most of the distance measures in the literature [1] consider the overlap or intersection between two histograms as a function of the distance value but do not take into account the similarity in the non-overlapping parts of the two histograms. For this reason, Rubner presented in [2] a new definition of the distance measure between histograms that overcomes this problem of non-overlapping parts. Called Earth Mover's Distance, it is defined as the minimum amount of work that must be performed to transform one histogram into another by moving distribution mass. This author used the simplex algorithm. Later, Cha presented in [1] three algorithms for obtaining the distance between one-dimensional histograms that use the Earth Mover's Distance. These algorithms compute the distance between histograms when the type of measurements are *nominal*, *ordinal* and *modulo* in $O(z)$, $O(z)$ and $O(z^2)$, respectively, where $z$ is the number of levels or bins.

Often, for specific set measurements, only a small fraction of the *bins* in a histogram contains significant information, i.e. most of the *bins* are empty. In such cases, the methods that use histograms as fixed-sized structures are not very efficient. For this reason, Rubner [2] presented variable-size descriptions called *signatures* that do not consider the empty bins.

In [3], the authors performed image retrieval based on colour histograms. Because the distance measure between colours is computationally expensive, they presented a low dimensional and easy-to-compute distance measure and showed that this was a lower boundary for the colour-histogram distance measure. An exact histogram-matching algorithm was presented in [4]. The aim of this algorithm was to study how various image characteristics affect colour reproduction by perturbing them in a known way.

In this paper we present the distances between histograms whose computational cost depends only on the non-empty bins rather than, as in the algorithms in [1,2], on the total number of bins. The type of measurements are *nominal*, *ordinal* and *modulo* and the computational cost is $O(z')$, $O(z')$ and $O(z'^2)$, respectively, where $z'$ is the number of non-empty bins in the histograms. In [5], we show that these distances are the same as the distances between the histograms in [1] but that the computational time for each comparison is lower when the histograms are large or sparse. The algorithms to compute them are not shown here due to lack of space (see [5]).

## 2. Histograms & Signatures

In this section, we formally define histograms and signatures. We end this section with a simple example to show the representations of the histograms and signatures given a set of measurements.

### 2.1. Histogram definition

Let $x$ be a measurement that can have one of $T$ values contained in the set $X=\{x_1,...x_T\}$. Consider a set of $n$ elements whose measurements of the value of $x$ are $A=\{a_1,...a_n\}$, where $a_t \in X$.

The histogram of the set $A$ along measurement $x$ is $H(x,A)$, which is an ordered list consisting of the number of occurrences of the discrete values of $x$ among the $a_t$. As we are interested only in comparing the histograms and sets of the same measurement $x$, $H(A)$ will be used instead of $H(x,A)$ without loss of generality. If $H_i(A)$, $1 \le i \le T$, denotes the number of elements of $A$ that have value $x_i$, then $H(A)=[H_1(A), \ldots, H_T(A)]$ where

$$H_i(A) = \sum_{t=1}^{n} C_{i,t}^{A} \quad \text{and} \quad C_{i,t}^{A} = \begin{cases} 1 & \text{if } a_t = x_i \\ 0 & \text{otherwise} \end{cases} \qquad (1)$$

The elements $H_i(A)$ are usually called *bins* of the histogram.

### 2.2. Signature definition

Let $H(A)=[H_1(A), \ldots, H_T(A)]$ and $S(A)=[S_1(A), \ldots, S_z(A)]$ be the histogram and the signature of the set *A*, respectively. Each $S_k(A)$, $1 \le k \le z \le T$ comprises a pair of terms, $S_k(A)=\{w_k, m_k\}$. The first term, $w_k$, shows the relation between the signature *S(A)* and the histogram *H(A)*. Therefore, if the $w_k=i$ then the second term, $m_k$, is the number of elements of *A* that have value $x_i$, i.e. $m_k=H_i(A)$ where $w_k<w_t \Leftrightarrow k<t$ and $m_k>0$.
The signature of a set is a lossless representation of its histogram in which the *bins* of the histogram whose value is 0 are not expressed implicitly. From the signature definition, we obtain the following expression,

$$H_{w_k}(A) = m_k \quad where \quad 1 \le k \le z \qquad (2)$$

### 2.3. Extended Signature

The **extended signature** is one in which some empty bins have been added. That is, we allow $m_i=0$ for some bins. This is a useful structure for ensuring that, given a pair of signatures to be compared, the number of bins is the same and that each bin in both signatures represents the same bin in the histograms.

### 2.4. Example

Figure 1 shows the histograms that represent sets *A* and *B* of 10 elements between 1 and 8.

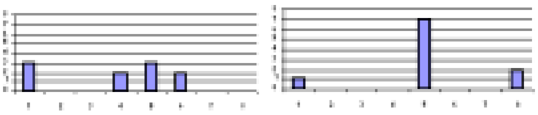

**Figure 1.** Histograms of sets *A* and *B*.

Figure 2 shows the signature representation of sets *A* and *B*. The length of the signatures is 4 and 3, respectively. Set A has 2 elements with a value of 6 since this value is represented by the bin 4 ($W_4^A=6$ ) and the value of the vertical axis is 2 at bin 4.

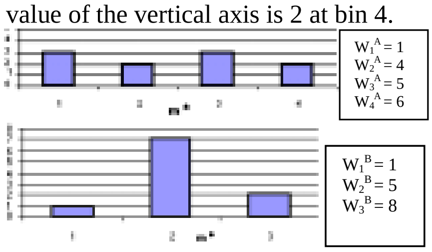


**Figure 2**. Signature representation of the sets *A* and *B*.

Figure 3 shows the extended signatures of the sets *A* and *B* with *5* bins. Note that the value that the extended signatures represents for each bin, $w_i$, is the same for both signatures.

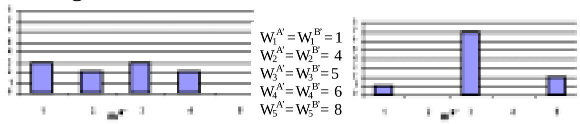


**Figure 3**. Extended Signatures *A'* and *B'*. The number of elements $m_i$ is represented graphically and the value of its elements is represented by $w_i$.

## 3. Type of measurement

We consider 3 types of measurements, called nominal (each value of the measurement is a name and there is no relation, such as greater than or lower than, between them), ordinal (the values are ordered) and modulo (the values are ordered but they form a ring because of the arithmetic modulo operation).
Corresponding to these three types of measurements, we define three measures of difference between two measurement levels $a \in X$ and $b \in X$, as follows:

**a) Nominal distance:**

$$d_{nom}(a,b) = \begin{cases} 0 & if \quad a=b \\ 1 & otherwise \end{cases} \qquad (3)$$

The distance value is either match or mismatch, which are mathematically represented by 0 or 1.

**b) Ordinal distance:**

$$d_{ord}(a,b) = |a-b| \qquad (4)$$

The distance value is computed by the absolute difference of each element.

**c) Modulo distance:**

$$d_{\mathrm{mod}}(a,b) = \begin{cases} |a-b| & if \ |a-b| \le T/2 \\ T-|a-b| & otherwise \end{cases} \qquad (5)$$

The distance value between two modulo measurement values is the interior difference of each element.

## 4. Distance between Signatures

For the following definitions of the distances, we assume that the extended signatures of *S(A)* and *S(B)* are *S(A')* and *S(B')*, where $S_i(A') = \{w_i^{A'}, m_i^{A'}\}$ and $S_i(B') = \{w_i^{B'}, m_i^{B'}\}$. The number of bins of *S(A)* and *S(B)* is $z^A$ and $z^B$ and the number of bins of both extended signatures is *z'*.

### 4.1. Nominal Distance

The nominal distance between the histograms in [3] is the number of elements that do not overlap or intersect. We redefine this distance using signatures as follows,

$$D_{nom}(S(A), S(B)) = \sum_{i=1}^{z'} \left| m_i^{A'} - m_i^{B'} \right| \qquad (6)$$

### 4.2. Ordinal Distance

The ordinal distance between two histograms was presented in [4] as the minimum work needed to transform one histogram into another. *H(A)* can be transformed into *H(B)* by moving elements to the left or to the right and the total number of all the necessary minimum movements is the distance between them. There are two operations. Suppose an element *a* that belongs to bin *i*. One operation is *move left (a)*. This result of this operation is that element *a* belongs to bin *i-1* and its cost is 1. This operation is impossible for the elements that belong to bin *1*. Another operation is *move right (a)*. Similarly, after this operation, *a* belongs to bin *i+1* and the cost is 1. The same restriction applies to the right-most bin. These operations are graphically represented by right-to-left arrows and left-to-right arrows. The total number of arrows is the distance value. This is the shortest movement to transform an histogram to the other.
The distance between signatures is defined as follows,

$$D_{ord}\left(S(A),S(B)\right)=\sum_{i=1}^{z'-1}\left[\left(w_{i+1}^{A'}-w_i^{A'}\right)\left|\sum_{j=1}^{i}\left(m_j^{A'}-m_j^{B'}\right)\right|\right] \quad (7)$$

The arrows do not have a constant size (or constant cost) but depend on the distance between bins. If element *a* belongs to bin *i*, the result of operation *move left (a)* is that the element *a* belongs to bin *i-1* and its cost is $w_i - w_{i-1}$. Similarly, after the operation *move right(a)*, the element *a* belongs to bin *i+1* and the cost is $w_{i+1} - w_i$. In equation (7), the number of arrows that go from bin *i* to bin *i+1* is described by the inner addition and the cost of these arrows is $w_{i+1} - w_i$.

### 4.3. Modulo Distance

One major difference in modulo type histograms or signatures is that the first bin and the last bin are considered to be adjacent to each other. It therefore forms a closed circle due to the nature of the data type. Transforming a modulo type histogram or signature into another while computing their distance should allow cells to move from the first bin to the last bin, or vice versa, at the cost of a single movement. Now, cells or blocks of earth can move from the first bin to the last bin with the operation *move left (1)* in the histogram case or *move left* $(w_1)$ in the signature case. Similarly, blocks can move from the last bin to the first one with the operations *move right (T)* in the histogram case or *move right* $(w_{z'})$ in the signature case.
The cost of these operations is calculated as the cost of the operations in the ordinal distance except for the movements of blocks from the first bin to the last or vice versa. For the distance between histograms, the cost, as in all the movements, is one. For the distance between signatures, the real distance between bins or the length of the arrows has to be considered. The cost of these movements is therefore the sum of three terms (figure 4): (a) the cost from the last bin of the signature, $w_{z'}$, to the last bin of the histogram, *T*; (b) the cost from the last bin of the histogram, *T*, to the first bin of the histogram, *1*; (c) the cost from the first bin of the histogram, *1*, to the first bin of the signature, $w_1$. The costs are then calculated as the length of these terms. The cost of (a) is $T-w_{z'}$, the cost of (b) is *1* (similar to the cost between histograms) and the cost of (c) is $w_1-1$. Therefore, the final cost from the last bin to the first or vice versa between signatures is $w_1-w_{z'}+T$.

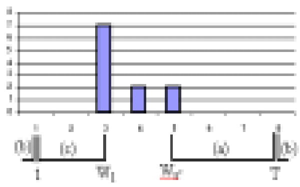

**Figure 4.** The three terms that need to be considered in order to compute the cost of moving blocks from the last bin to the first or vice versa.

Due to the previously explained modulo properties, we can transform one signature or histogram into another in several ways. In one of these ways, there is a minimum distance whose number of movements (or the cost of the arrows and the number of arrows) is the lowest. If there is a borderline between bins that has both directional arrows, they are cancelled out. These movements are redundant, so the distance cannot be obtained through this configuration of arrows. To find the minimum configuration of arrows, we can add a complete chain in the histogram or signature of the same directional arrows and the opposite arrows on the same border between bins are then cancelled out. The modulo distance between signatures is defined as

$$D_{\mathrm{mod}}\left(S(A),S(B)\right)= \min_c\left\{\sum_{i=1}^{z'-1}\left[\left(w_{i+1}^{A'}-w_i^{A'}\right)\left|c+\sum_{j=1}^{i}\left(m_j^{A'}-m_j^{B'}\right)\right|\right]+\left(w_1^{A'}-w_{z'}^{A'}+T\right)|c|\right\} \quad (8)$$

The cost of moving a block of earth from one bin to another is not 1 but the length of the arrows or the distance between the bins (as explained in the ordinal distance between signatures). The cost of the movement of blocks from the first bin to the last or vice versa is $w_1-w_{z'}+T$ and the cost of the other movements is $w^{A'}_{i+1}-w^{A'}_{i}$. The term *c* represents the chains of left arrows or right arrows added to the current arrow representation. The absolute value of *c* at the end of the expression is the number of chains added to the current representation. It comes from the cost of the arrows from the last bin to the first or vice versa.

**Example.** Figure 5 shows five different transformations of signature *S(A)* to signature *S(B)* and their related costs. In the first transformation, one chain of right arrows is added (c=1). In the second transformation, no chains are added (c=0), so the cost is the same as the ordinal distance. In the third to the last transformations, 1, 2 and 3 chains of left arrows are added, respectively. We can see that the minimum cost is 6 and c=-2, the distance value is 6 for the modulo distance and 14 for the ordinal distance.

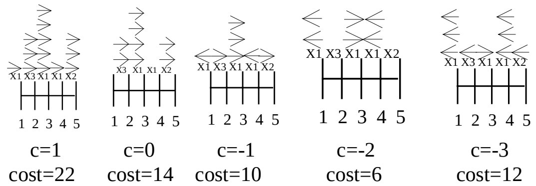


**Figure 5.** Five different transformations of signature *S(A)* to signature *S(B)* with their related *c* and cost obtained.

## 5. Robot Positioning by Image Retrieval

Our method has been tested using indoor scenes. These scenes were used for robot positioning. In the learning stage, the robot is guided through the offices and corridors while the exact position is introduced and the robot captures the scenes. In the recognition stage, the robot assumes to be in the position that captures the most similar scene obtained in the learning stage [6]. In [7], a same robot-positioning method was used. The main difference is that they used structural information of the image and they needed to segment the image. For this reason, the ratio of recognition was supposed to be higher but also the computational time. The main advantage of our method is that the image has not to be processed; only the signature of it is needed.

Figure 6 represents a possible scene; we have taken 3 images with a slight difference of the position. We have used the histograms of the luminance to test the cardinal and ordinal distance and the histogram of the hue to test the modulo distance.

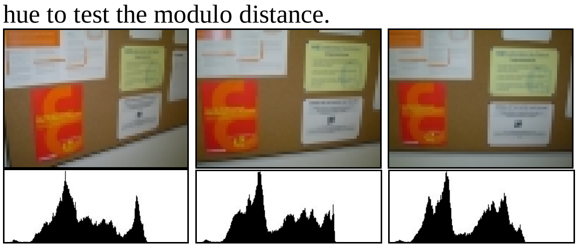

**Figure 6.** Three slightly different images from the same scene and their luminance histograms

Table 1 shows the run time and ratio of recognition obtained from three experiments. The first one (a), the results where computed using the cardinal distance on the luminance histograms. The second one (b), the cardinal distance was changed by the ordinal distance. And the third experiment (c), the modulo distance was computed on the hue histograms. From each experiment, we obtained the run time and the ratio of recognition in 3 cases. 1: histograms. 2: signatures. 3: filtered signatures. In this case, the bins that had fewer elements than a threshold were removed. The threshold of the filter was situated as much higher as possible, when the ratio of recognition began to decrease. In all the cases, the run time was normalized such that the run time of the histogram in the first case was 100.

It is interesting to realize the decrease on the run time in the case of the modulo distance when the filter is applied. There is a decrease from 526 to 78.

| | **Card Dist** | | | **Ord Dist** | | | **Mod Dist** | | |
|---|---|---|---|---|---|---|---|---|---|
| | Histo. | Sign. | | Histo. | Sign. | | Histo. | Sign. | |
| | | No filter | Filter | | No filter | Filter | | No filter | Filter |
| Run time | 100 | 85 | 51 | 105 | 87 | 45 | 526 | 233 | 78 |
| %Recog. | 70.8 | 70.8 | 70.5 | 87.5 | 87.5 | 87.3 | 91.6 | 91.6 | 91.4 |

(a) Luminance (b) Luminance (c) Hue

**Table 1**. Run time and ratio of recognition obtained from three experiments on the luminance histograms (a) and (b) and on the hue histograms (c).

## 6. References


[1] S.-H. Cha, S. N. Srihari, "On measuring the distance between histograms" *Pattern Recognition* 35, pp: 1355–1370, 2002.

[2] Y. Rubner, C. Tomasi, and L. J. Guibas, "A Metric for Distributions with Applications to Image Databases" *Inte.l Journal of Comput Vision* 40(2),pp: 99-121,2000.

[3] J.Hafner, J.S. Sawhney, W. Equitz, M. Flicker & W. Niblack, "Efficient Colour Histogram Indexing for Quadratic Form Distance Functions", *Trans. On Pattern Analysis and Machine Intell.*, 17 (7), pp:729-735, 1995.

[4] J. Morovic, J. Shaw & P-L. Sun, "A fast, non-iterative and exact histogram matching algorithm", *Pattern Recognition Letters* 23, pp:127–135, 2002.

[5] F. Serratosa & A. Sanfeliu, "Signatures versus Histograms: Definitions, Distances and Algorithms", *Pattern recognition*, 2006.

[6] G. Wells, Ch. Venaille & C. Torras, "Vision-based robot positioning using neural networks", *Image & Vision Computing*,Vol14(10),1996,pp:715-732.

[7] E.Staffeti, A.Grau, F.Serratosa & A.Sanfeliu, "Object and Image indexing based on Region Connection Calculus and Oriented Matroids theory", *Discrete Applied Mathematics*, Vol. 147, (2-3),2005 pp:345-361.