



**HAL**  
open science

## Transportation Distances on the Circle

Julien Rabin, Julie Delon, Yann Gousseau

► **To cite this version:**

Julien Rabin, Julie Delon, Yann Gousseau. Transportation Distances on the Circle. 2009. hal-00399832v2

**HAL Id: hal-00399832**

**<https://hal.science/hal-00399832v2>**

Preprint submitted on 6 Jul 2009 (v2), last revised 1 Jun 2010 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# TRANSPORTATION DISTANCES ON THE CIRCLE

JULIEN RABIN, JULIE DELON, AND YANN GOUSSEAU

ABSTRACT. In this contribution, we study Monge-Kantorovich distances between discrete set of points on the unit circle  $S^1$ , when the ground distance between two points  $x$  and  $y$  on the circle is defined as  $c(x, y) = \min(|x - y|, 1 - |x - y|)$ . We first prove that computing a Monge-Kantorovich distance between two given sets of pairwise different points boils down to cut the circle at a well chosen point and to compute the same distance on the real line.

This result is then used to prove a formula on the Earth Mover's Distance [3], which is a particular Monge-Kantorovich distance. This formula asserts that the Earth Mover's Distance between two discrete circular normalized histograms  $f = (f[i])_{i=0, \dots, N-1}$  and  $g = (g[i])_{i=0, \dots, N-1}$  on  $N$  bins can be computed by

$$(1) \quad \text{CEMD}(f, g) = \min_{k \in \{0, \dots, N-1\}} \|F_k - G_k\|_1,$$

where  $F_k$  and  $G_k$  are the cumulative histograms of  $f$  and  $g$  starting at the  $k^{\text{th}}$  quantization bin. This formula is used in recent papers [1, 2] on the matching of local features between images, where the Earth Mover's Distance is used to compare circular histograms of gradient orientations.

## 1. INTRODUCTION

Consider two discrete set of points  $\{x_1, \dots, x_P\}$  and  $\{y_1, \dots, y_P\}$  on the unit circle  $S^1$ , and the corresponding discrete distributions

$$f = \frac{1}{P} \sum_{k=1}^P \delta_{x_k}, \text{ and } g = \frac{1}{P} \sum_{k=1}^P \delta_{y_k},$$

where the notations  $x_k, y_k$  are used equally for points on the unit circle or for their coordinates in  $[0, 1[$ . For any  $\lambda > 0$ , the  $\lambda$ -Monge-Kantorovich distance between  $f$  and  $g$  is defined as

$$(2) \quad \text{MK}_\lambda(f, g) := \min_{(\alpha_{i,j}) \in \mathcal{M}} \left( \sum_{i=1}^N \sum_{j=1}^N \alpha_{i,j} c(x_i, y_j)^\lambda \right)^{\frac{1}{\lambda}}, \text{ where}$$

$$(3) \quad \mathcal{M} = \{(\alpha_{i,j}); \alpha_{i,j} \geq 0, \sum_j \alpha_{i,j} = \frac{1}{P}, \sum_i \alpha_{i,j} = \frac{1}{P}\}$$

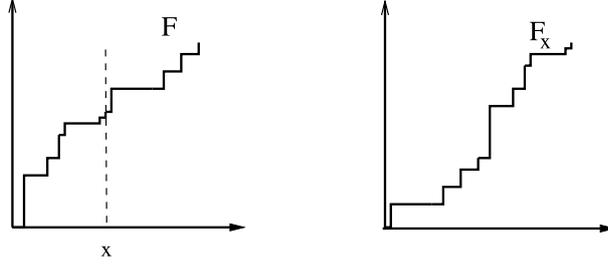
and where  $c(\cdot, \cdot)$  is the periodic distance on  $[0, 1[$

$$(4) \quad c(x, y) = \min(|x - y|, 1 - |x - y|).$$

The Earth Mover's Distance [3] is the distance obtained when  $\lambda = 1$ .

Now, it is well known (see for example the introduction of [4]) that the  $\lambda$ -Monge-Kantorovich distance between  $f$  and  $g$  equals

$$(5) \quad \text{MK}_\lambda(f, g) := \min_{\sigma \in \Sigma_P} W_\sigma^\lambda(f, g)$$

FIGURE 1.  $F$  on the left and  $F_x$  on the right.

where  $\Sigma_P$  is the set of permutations of  $\{1, \dots, P\}$  and where

$$(6) \quad W_\sigma^\lambda(f, g) := \frac{1}{P} \left( \sum_k c(x_k, y_{\sigma(k)})^\lambda \right)^{\frac{1}{\lambda}}$$

is the cost of transporting  $f$  to  $g$  with the permutation  $\sigma$ . In other words, finding the optimal transportation between  $f$  and  $g$  boils down to find the optimal permutation  $\sigma$  between the points  $\{x_k\}$  and  $\{y_k\}$ .

**1.1. Paths.** If  $x$  and  $y$  are two different points of  $S^1$ , we consider  $\gamma(x, y)$  **the geodesic path linking  $x$  and  $y$  on  $S^1$**  (the path is supposed open : it does not contain  $x$  and  $y$ ). This path is always unique except in the case where  $x$  and  $y$  are in opposite positions on the circle. In this case, we choose  $\gamma(x, y)$  as the path going from  $x$  to  $y$  in the trigonometric direction. A path  $\gamma(x, y)$  is said to be **positive** if it goes from  $x$  to  $y$  in the trigonometric direction. If the path goes from  $x$  to  $y$  in the opposite direction, it is said to be **negative**.

**1.2. Cumulative distribution functions.** We define the cumulative distribution function of  $f = \frac{1}{P} \sum_{k=1}^P \delta_{x_k}$  on  $[0, 1[$  as

$$(7) \quad \forall y \in [0, 1[, \quad F(y) = \frac{1}{P} \sum_{k=1}^P \mathbb{1}_{\{x_k \in [0, y]\}}.$$

Observe that on  $[0, 1[$  seen as a the unit circle  $S^1$ , no strict order can be defined between points. We can define as many cumulative distribution functions as there are starting points on the circle. If  $x$  is a point in  $[0, 1[$ , the  $x$ -cumulative distribution function  $F_x$  of  $f$  can be defined by choosing  $x$  as the reference point on the circle  $S^1$  and by cumulating the mass in the trigonometric order from this new reference point. More precisely,  $F_x$  can be computed as

$$(8) \quad \forall y \in [0, 1[, \quad F_x(y) = \begin{cases} \frac{1}{P} \sum_{k=1}^P \mathbb{1}_{\{x_k \in [x, y+x]\}} & \text{if } y < 1 - x, \\ \frac{1}{P} \sum_{k=1}^P \mathbb{1}_{\{x_k \in [x, 1] \cup [0, y-1+x]\}} & \text{if } y \geq 1 - x. \end{cases}$$

Observe that  $F_x$  is an increasing function and that for each  $(x, y)$  in  $[0, 1]^2$ ,  $F_x(y) = F(y + x - 1) + 1 - F(x)$  if  $y \geq 1 - x$ , and  $F_x(y) = F(y + x) - F(x)$  if  $y < 1 - x$ . An example of a cumulative distribution  $F$  and its corresponding  $x$ -cumulative distribution  $F_x$  is shown on Figure 1.

## 2. COMPUTING THE MONGE-KANTOROVICH DISTANCE ON THE CIRCLE

In the following, we show that the Monge-Kantorovich distance on the circle can be computed very easily for  $\lambda = 1$  (this case corresponds to the EMD distance). In short, we show that if  $\sigma$  is an optimal permutation for (5), there is always a point on the circle which is not contained in any optimal path of  $\sigma$ . This result is proven first for  $\lambda > 1$  and for any optimal permutation  $\sigma$ , then for  $\lambda = 1$  and a well chosen optimal permutation. The main consequence of this result is that the computation of the Monge-Kantorovich distance on the circle boils down to the computation of the same distance on an interval of  $\mathbb{R}$ .

**Proposition 1.** *Assume that  $\lambda > 1$ . Let  $x_1, \dots, x_P$  and  $y_1, \dots, y_P$  be  $P$  points in  $[0, 1[$ . Assume that all these points are pairwise different. Then for each permutation  $\sigma$  of  $\Sigma_P$  which minimizes (5) with the cost (4), there exists  $k \in \{1, \dots, P\}$  such that for all  $l \neq k$ ,  $x_k \notin \gamma(x_l, y_{\sigma(l)})$ .*

The proof of this proposition needs the following lemma, which describes some properties of the geodesic paths  $\gamma(x_l, y_{\sigma(l)})$  obtained when  $\sigma$  is a minimizer of (5) with  $\lambda > 1$ .

**Lemma 1.** *Assume that  $\lambda > 1$ . Let  $\sigma$  be a minimizer of (5) and let  $\gamma_l = \gamma(x_l, y_{\sigma(l)})$  and  $\gamma_k = \gamma(x_k, y_{\sigma(k)})$  (with  $l \neq k$ ) be two geodesic paths for the assignment defined by  $\sigma$ . Assume also that  $x_l \neq x_k$  and  $y_{\sigma(l)} \neq y_{\sigma(k)}$ . Then, one of the following holds:*

- $\gamma_l \cap \gamma_k = \emptyset$ ;
- $\gamma_l \cap \gamma_k \neq \emptyset$  and in this case  $\gamma_l$  and  $\gamma_k$  have the same direction (both positive or both negative) and neither of them is contained in the other.

*Proof.* Assume that  $\gamma_l \cap \gamma_k \neq \emptyset$ . If  $\gamma_l \cap \gamma_k$  is equal to  $\gamma(x_l, x_k)$  or  $\gamma(y_{\sigma(l)}, y_{\sigma(k)})$ , then

$$c(x_l, y_{\sigma(l)})^\lambda + c(x_k, y_{\sigma(k)})^\lambda > c(x_l, y_{\sigma(k)})^\lambda + c(x_k, y_{\sigma(l)})^\lambda,$$

which contradicts the optimality of  $\sigma$ . Moreover, if for example the path  $\gamma_l$  is included in  $\gamma_k$ , then the strict convexity of the function  $x \mapsto |x|^\lambda$  (for  $\lambda > 1$ ) implies

$$c(x_l, y_{\sigma(l)})^\lambda + c(x_k, y_{\sigma(k)})^\lambda > c(x_l, y_{\sigma(k)})^\lambda + c(x_k, y_{\sigma(l)})^\lambda,$$

which also contradicts the optimality of  $\sigma$ . Thus,  $\gamma_l \cap \gamma_k$  is equal to  $\gamma(x_l, y_{\sigma(k)})$  or to  $\gamma(x_k, y_{\sigma(l)})$  and it follows that  $\gamma_k$  and  $\gamma_l$  are either both positive or both negative.  $\square$

**Proof of Proposition 1** Let  $\sigma$  be a minimizer of (5). In the following, we will denote by  $\gamma_l$  the geodesic path  $\gamma(x_l, y_{\sigma(l)})$ . We can assume without loss of generality that the points  $x_1, \dots, x_P$  are in trigonometric order on the circle.

Assume that for each  $l \in \{1, \dots, P\}$ , there exists  $q(l) \neq l$  such that  $x_l$  belongs to the open path  $\gamma_{q(l)}$ . Then, for each  $l$ , we have  $\gamma_{q(l)} \cap \gamma_l \neq \emptyset$ , which means that the geodesic paths  $\gamma_{q(l)}$  and  $\gamma_l$  are either both positive or both negative (from lemma 1). Assume for instance that they are both positive and let us show that in this case  $x_l \in \gamma_{l-1}$  (with  $l-1 = P$  if  $l = 0$ ). If  $q(l) = l-1$ , there is nothing to prove. If  $q(l) \neq l-1$ , it means in particular that  $x_{q(l)}, x_{l-1}, x_l$  are in trigonometric order on the circle. Since  $\gamma_{q(l)}$  is a positive path starting from  $x_{q(l)}$  and containing  $x_l$ , it follows that  $\gamma_{q(l)}$  contains  $x_{l-1}$  (recall that the points are assumed to be pairwise different, in particular  $x_{l-1} \neq x_{q(l)}$ ). Thus  $\gamma_{l-1} \cap \gamma_{q(l)} \neq \emptyset$ , which implies that  $\gamma_{l-1}$  is positive. Now,  $x_l$  must be in  $\gamma_{l-1}$ , otherwise we would have  $\gamma_{l-1} \subset \gamma_{q(l)}$ , which is impossible from lemma 1. Thus, if the paths  $\gamma_{q(l)}$  and  $\gamma_l$  are both positive,  $x_l \in \gamma_{l-1}$ .

In the same way, if  $\gamma_{q(l)}$  and  $\gamma_l$  are both negative, then  $x_l \in \gamma_{l+1}$ . In any case, for each  $l \in \{1, \dots, P\}$ ,  $x_l \in \gamma_{l-1} \cup \gamma_{l+1}$  (with the obvious convention  $\gamma_{P+1} = \gamma_1, \gamma_0 = \gamma_P$ ).

Now, suppose that for a given  $k \in \{1, \dots, P\}$ ,  $x_k$  is in  $\gamma_{k-1}$ . Then,  $\gamma_{k-1}$  and  $\gamma_k$  have the same direction. From lemma 1, it follows that  $x_{k-1}$  cannot be contained in  $\gamma_k$ . Since we know that  $x_{k-1} \in \gamma_{k-2} \cup \gamma_k$ ,  $x_{k-1}$  must be in  $\gamma_{k-2}$ . Recursively, for each  $l \in \{1, \dots, P\}$ ,  $x_l \in \gamma_{l-1}$ . It follows that for each  $l \in \{1, \dots, P\}$ ,  $c(x_l, y_{\sigma(l-1)}) < c(x_{l-1}, y_{\sigma(l-1)})$ , and then

$$(9) \quad \sum_{l=1}^P c(x_l, y_{\sigma(l)})^\lambda > \sum_{l=1}^P c(x_{l+1}, y_{\sigma(l)})^\lambda,$$

which contradicts the fact that  $\sigma$  is a minimizer of (5). We come to the same conclusion if for a given  $k \in \{1, \dots, P\}$ ,  $x_k$  is in  $\gamma_{k+1}$   $\square$

The same result can be proven for  $\lambda = 1$  with the difference that it is only satisfied for a good choice of the permutation  $\sigma$  which minimizes (5), and not for all of these permutations. This result can be seen as a limit version of proposition 1 when  $\lambda \rightarrow 1$ .

**Corollary 1.** *Assume that  $\lambda = 1$ . Let  $x_1, \dots, x_P$  and  $y_1, \dots, y_P$  be  $P$  points in  $[0, 1]$ . Assume that all these points are pairwise different. Then there exists a permutation  $\sigma$  of  $\Sigma_P$  which minimizes (5) and a point  $x_k \in \{x_1, \dots, x_P\}$  such that for all  $l \neq k$ ,  $x_k \notin \gamma(x_l, y_{\sigma(l)})$ .*

*Proof.* We know that for any  $\lambda > 1$ , if  $\sigma_\lambda$  minimizes the cost  $\sigma \mapsto W_\sigma^\lambda(f, g)$ , there exists  $k \in \{1, \dots, P\}$  such that for all  $l \neq k$ ,  $x_k \notin \gamma_l = \gamma(x_l, y_{\sigma_\lambda(l)})$ .

Now, observe that if  $\sigma$  and the points  $x_1, \dots, x_P, y_1, \dots, y_P$  are fixed,  $W_\sigma^\lambda(f, g)$  is a continuous function of  $\lambda$ . Thus, for each  $\varepsilon > 0$ , there exists  $\beta > 1$ , such that for all  $\lambda \in [1, \beta]$ ,  $|W_\sigma^\lambda(f, g) - W_\sigma^1(f, g)| \leq \varepsilon$ . Since  $\Sigma_P$  is a finite set, we can chose  $\beta$  close enough to 1 such that this property holds for every  $\sigma$  in  $\Sigma_P$ . We can also chose  $\beta$  such that  $|\min_\sigma W_\sigma^1(f, g) - \min_\sigma W_\sigma^\lambda(f, g)| \leq \varepsilon$  (the minimum of a finite set of continuous functions is a continuous function). It follows that for  $\lambda \in [1, \beta]$ ,

$$\begin{aligned} |\min_\sigma W_\sigma^1(f, g) - W_{\sigma_\lambda}^1(f, g)| &\leq |\min_\sigma W_\sigma^1(f, g) - \min_\sigma W_\sigma^\lambda(f, g)| \\ &\quad + |W_{\sigma_\lambda}^\lambda(f, g) - W_{\sigma_\lambda}^1(f, g)| \\ &\leq 2\varepsilon. \end{aligned}$$

Thus, when  $\lambda$  is close enough to 1, a minimizer  $\sigma_\lambda$  of  $W_\sigma^\lambda(f, g)$  is also a minimizer of  $W_\sigma^1(f, g)$ . Which proves that there exists at least one minimizer  $\sigma$  of  $\sigma \mapsto W_\sigma^1(f, g)$  such that  $x_k \notin \gamma(x_l, y_{\sigma(l)})$  for some  $k \in \{1, \dots, P\}$  and all  $l \neq k$ .  $\square$

We can now draw a first conclusion on the form of the Monge Kantorovich distance when all points are pairwise different.

**Corollary 2.** *Assume that  $x_1, \dots, x_P$  and  $y_1, \dots, y_P$  are pairwise different. Then,*

$$(10) \quad \forall \lambda \geq 1, \quad \text{MK}_\lambda(f, g) = \left( \inf_{x \in S^1} \int |F_x^{-1} - G_x^{-1}|^\lambda \right)^{1/\lambda},$$

where  $F_x^{-1}$  and  $G_x^{-1}$  are the pseudo-inverses of the increasing functions  $F_x$  and  $G_x$ , defined as  $F_x^{-1}(y) = \inf\{t; F_x(t) > y\}$  and  $G_x^{-1}(y) = \inf\{t; G_x(t) > y\}$ .

*Proof.* Proposition 1 and the above corollary show that for any  $\lambda \geq 1$ , if  $x_1, \dots, x_P$  and  $y_1, \dots, y_P$  are pairwise different, we can choose some optimal permutation  $\sigma$  for which there is some point  $x_k$  which is not contained in any path (recall that paths are defined

as open: they do not contain their boundaries). Since all points are supposed pairwise different, the only path meeting all the neighborhoods of  $x_k$  is  $\gamma_k$ . It follows that there exists some open set on one side of  $x_k$  and not containing  $x_k$  which does not cross any path of the optimal permutation  $\sigma$ . The middle  $x$  of this open set is not contained in any path of  $\sigma$ . We can thus cut the circle  $S^1$  at  $x$  and reduce the transportation problem on the circle to the transportation problem on the real line. The optimal permutation  $\sigma$  is thus given by the sorting of the points (formula (72) in [4]), taking  $x$  as the reference point on the circle. This means that when points are pairwise different, we have

$$(11) \quad \forall \lambda \geq 1, \quad \text{MK}_\lambda(f, g) = \left( \inf_{x \in S^1} \int |F_x^{-1} - G_x^{-1}|^\lambda \right)^{1/\lambda},$$

where  $F_x^{-1}$  and  $G_x^{-1}$  are the pseudo-inverses of the increasing functions  $F_x$  and  $G_x$ .  $\square$

Observe that in the case where  $\lambda = 1$ , this result can be rewritten

$$(12) \quad \text{MK}_1(f, g) = \inf_{x \in S^1} \|F_x - G_x\|_1.$$

In order to generalize this result to the case where the points  $x_1, \dots, x_P$  and  $y_1, \dots, y_P$  can coincide, we first show that  $\text{MK}_1(f, g)$  can also be written  $\min_{l \in \{-P, \dots, P\}} \|F - G - \frac{l}{P}\|_1$ , which will be easier to handle. Obviously, the quantity  $\|F_x - G_x\|_1$  can be written as

$$\begin{aligned} \|F_x - G_x\|_1 &= \int_0^{1-x} |F(t+x) - F(x) - G(t+x) + G(x)| dt \\ &\quad + \int_{1-x}^1 |F(t+x-1) - F(x) - G(t+x-1) + G(x)| dt \\ &= \int_x^1 |F(t) - F(x) - G(t) + G(x)| dt + \int_0^x |F(t) - F(x) - G(t) + G(x)| dt \\ &= \|F - F(x) - G + G(x)\|_1. \end{aligned}$$

Now, observe that  $F(x) - G(x)$  takes its values in  $\{\frac{l}{P}; l = -P, \dots, P\}$ . Moreover, since  $F - G$  takes only a finite number of values, it is easy to see that  $\|F - G - \frac{l}{P}\|_1$  reaches its minimum at one (non necessarily unique) of these values. Thus,

$$\text{MK}_1(f, g) = \min_{l \in \{-P, \dots, P\}} \|F - G - \frac{l}{P}\|_1.$$

In the following corollary, we show that this formula can be generalized to the case where some of the points  $x_1, \dots, x_P$  and  $y_1, \dots, y_P$  coincide.

**Corollary 3.** *Let  $f = \frac{1}{P} \sum_{k=1}^P \delta_{x_k}$  and  $g = \frac{1}{P} \sum_{k=1}^P \delta_{y_k}$  be two discrete distributions on  $S^1$  (some of the points  $x_1, \dots, x_P$  and  $y_1, \dots, y_P$  can coincide). Then, the 1-Monge Kantorovich distance between  $f$  and  $g$  can be computed as*

$$\text{MK}_1(f, g) = \min_{l \in \{-P, \dots, P\}} \|F - G - \frac{l}{P}\|_1.$$

*Proof.* Let  $x_1, \dots, x_P, y_1, \dots, y_P$  be  $2P$  points on  $S^1$ , not necessarily different. For every  $\varepsilon > 0$ , we can build  $x_1^\varepsilon, \dots, x_P^\varepsilon, y_1^\varepsilon, \dots, y_P^\varepsilon$ , all pairwise different, such that  $\forall k \in \{1, \dots, P\}$ ,  $c(x_k^\varepsilon, x_k) \leq \varepsilon$  and  $c(y_k^\varepsilon, y_k) \leq \varepsilon$ . Let  $f^\varepsilon = \frac{1}{P} \sum_{k=1}^P \delta_{x_k^\varepsilon}$  and  $g^\varepsilon = \frac{1}{P} \sum_{k=1}^P \delta_{y_k^\varepsilon}$

be the corresponding distributions. Then,

(13)

$$\forall \sigma \in \Sigma_P, \quad |W_\sigma^1(f^\varepsilon, g^\varepsilon) - W_\sigma^1(f, g)| \leq \sum_{k=1}^P |c(x_k^\varepsilon, y_{\sigma(k)}^\varepsilon) - c(x_k, y_{\sigma(k)})| \leq 2P\varepsilon = C\varepsilon.$$

Since  $\text{MK}_1$  is a minimum over a finite set, it follows that

$$\text{MK}_1(f^\varepsilon, g^\varepsilon) \xrightarrow{\varepsilon \rightarrow 0} \text{MK}_1(f, g).$$

Now, if we write  $F^\varepsilon$  and  $G^\varepsilon$  the cumulative distribution functions of the distributions  $f^\varepsilon$  and  $g^\varepsilon$ , then

$$\forall l \in \{-P, \dots, P\}, \quad \|F^\varepsilon - G^\varepsilon - \frac{l}{P}\|_1 \xrightarrow{\varepsilon \rightarrow 0} \|F - G - \frac{l}{P}\|_1.$$

Since  $l$  takes a finite number of values,

$$\min_{l \in \{-P, \dots, P\}} \|F^\varepsilon - G^\varepsilon - \frac{l}{P}\|_1 \xrightarrow{\varepsilon \rightarrow 0} \min_{l \in \{-P, \dots, P\}} \|F - G - \frac{l}{P}\|_1.$$

From these two convergences, we can conclude that

$$(14) \quad \text{MK}_1(f, g) = \min_{l \in \{-P, \dots, P\}} \|F - G - \frac{l}{P}\|_1.$$

□

Observe that an alternative proof of this last result when  $\lambda = 1$  was proposed by Werman et al. in [5], with some inaccuracies in the case where points may coincide.

Now, using the same sequence of equalities as before, and noting that  $\|F - G - \frac{l}{P}\|_1$  must reach its minimum at one of the values  $\|F - G - F(x) + G(x)\|_1$ , we can conclude that formula (12) is satisfied in the case where points may coincide.

### 3. CONSEQUENCES IN THE CASE OF DISCRETE HISTOGRAMS

Consider two discrete histograms  $f = (f[i])_{i=0 \dots N-1}$  and  $g = (g[i])_{i=0 \dots N-1}$ , sampled on  $N$  equal bins. We suppose that  $f$  and  $g$  are normalized, in the sense that  $\sum_{i=0}^{N-1} f[i] = \sum_{i=0}^{N-1} g[i] = 1$ . Now, assume that these histograms are **circular**, which means that the bins 0 and  $N - 1$  are neighbors. In this case, these two histograms can be considered as probability distributions on the unit circle  $S^1$ , or equivalently as periodic distributions of period 1 on  $\mathbb{R}$ . These distributions can then be written as

$$f = \sum_{i=0}^{N-1} f[i] \delta_{i/N}, \quad \text{and} \quad g = \sum_{k=0}^{N-1} g[k] \delta_{k/N},$$

Now, if all the weights  $f[i]$  and  $g[k]$  are rational, which is the case for numerical histograms, it is always possible to replicate these points as many times as necessary in order to write

$$f = \frac{1}{P} \sum_{k=1}^P \delta_{x_k}, \quad \text{and} \quad g = \frac{1}{P} \sum_{k=1}^P \delta_{y_k},$$

where  $\{x_1, \dots, x_P\}$  and  $\{y_1, \dots, y_P\}$  are two sets of points on the unit circle  $S^1$ . As a result, several of these points may coincide. For instance, if  $N = 2$  and  $f = (3/4, 1/4)$ , we can write  $f = \frac{1}{4}(\delta_0 + \delta_0 + \delta_0 + \delta_{\frac{1}{2}})$ .

The circular Earth Mover's Distance between the discrete histograms  $f$  and  $g$  can thus be computed as

$$(15) \quad \text{CEMD}(f, g) = \text{MK}_1(f, g) = \min_{l \in \{-P, \dots, P\}} \|F - G - \frac{l}{P}\|_1 = \min_{l \in \{-P, \dots, P\}} \sum_{k=-P}^P \left| \frac{k-l}{P} \right| \omega_k,$$

where  $\omega_k = \text{measure}\{z \in [0, 1[; F(z) - G(z) = \frac{k}{P}\}$ . In other words, the value  $\frac{l}{P}$  where this quantity is minimum can be computed as a weighted median of the values  $F(z) - G(z)$ . In the case of a discrete histograms, the points  $x_k$  and  $y_k$  are located on a regular grid of  $N$  bins. Consequently,

$$(16) \quad \text{MK}_1(f, g) = \frac{1}{N} \min_{j \in \{0, \dots, N-1\}} \sum_{i=0}^{N-1} |F[i] - G[i] - F[j] + G[j]| = \frac{1}{N} \sum_{i=0}^{N-1} |F[i] - G[i] - \mu|,$$

where  $\mu$  is the median of all the values  $F[i] - G[i]$ .

In the same way, for discrete histograms, formula (12) becomes

$$(17) \quad \text{CEMD}(f, g) = \min_{k \in \{0, \dots, N-1\}} \|F_k - G_k\|_1,$$

where  $F_k$  and  $G_k$  are the cumulative histograms of  $f$  and  $g$  starting at the  $k^{\text{th}}$  quantization bin. For each  $k$  in  $\{0, \dots, N-1\}$

$$F_k[i] = \begin{cases} \sum_{j=k}^{k+i} f[j] & \text{if } 0 \leq i \leq N-k-1 \\ \sum_{j=k}^{N-1} f[j] + \sum_{j=0}^{k+i-N} f[j] & \text{if } N-k \leq i \leq N-1 \end{cases}.$$

The definition is similar for  $G_k$  by replacing  $f$  by  $g$ . Note that  $F_k[i] = F[k+i] - F[k-1]$  if  $i \leq N-k-1$  and  $F_k[i] = F[k+i-N+1] + 1 - F[k-1]$  if  $i > N-k-1$ .

In other words, the distance  $\text{CEMD}(f, g)$  is also the minimum in  $k$  of the  $L^1$  distance between  $F_k$  and  $G_k$ , the cumulative histograms of  $f$  and  $g$  starting at the  $k^{\text{th}}$  quantization bin.

Observe that this formula remains valid for any translated version of  $F_k$  and  $G_k$ . In particular, if we define

$$\tilde{F}_k[i] = \begin{cases} \sum_{j=k}^i f[j] & \text{if } k \leq i \leq N-1 \\ \sum_{j=k}^{N-1} f[j] + \sum_{j=0}^i f[j] & \text{if } 0 \leq i \leq k-1 \end{cases},$$

the distance  $\text{CEMD}(f, g)$  can also be computed as

$$(18) \quad \text{CEMD}(f, g) = \min_{k \in \{0, \dots, N-1\}} \|\tilde{F}_k - \tilde{G}_k\|_1.$$

#### ACKNOWLEDGMENTS

Delon acknowledges the support of the French Agence Nationale de la Recherche (ANR), under grant BLAN07-2-183172, Optimal transport: Theory and applications to cosmological reconstruction and image processing (OTARIE).

## REFERENCES

- [1] Julien Rabin, Julie Delon, and Yann Gousseau. Circular earth mover's distance for the comparison of local features. In *Proc. ICPR*. IEEE Computer Society, 2008.
- [2] Julien Rabin, Julie Delon, and Yann Gousseau. A statistical approach to the matching of local features. *Submitted to SIAM Journal on Imaging Sciences*, 2009.
- [3] Yossi Rubner, Carlo Tomasi, and Leonidas J. Guibas. The Earth Mover's distance as a metric for image retrieval. *Int. J. Comput. Vision*, 40(2):99–121, 2000.
- [4] C. Villani. *Topics in optimal transportation*. American Math. Soc., 2003.
- [5] M. Werman, S. Peleg, R. Melter, and TY Kong. Bipartite graph matching for points on a line or a circle. *Journal of Algorithms*, 7(2):277–284, 1986.