A FRAMEWORK FOR STEREO VISION VIA OPTIMAL TRANSPORT

MATTIA GALEOTTI, ALESSANDRO SARTI, GIOVANNI CITTI

Abstract. We present a theoretical framework for a stereo vision method via optimal transport tools. We consider two aligned optical systems and we develop the matching between the two pictures line by line. By considering a regularized version of the optimal transport, we can speed up the computation and obtain a very accurate disparity function evaluation. Moreover, via this same method we can approach successfully the case of images with occluded regions.

1. Intro. Stereo vision is the ability of reconstructing a 3-D representation of a given situation from the data of two 2-D pictures taken with a small shift. This process is clearly fundamental in human vision, has been analyzed through a great variety of mathematical tools and it is central also in many technological applications. In particular stereo vision techniques employ two cameras with the same focal length looking at the same scene. The fundamental step in the stereo reconstruction consists in building a correspondence between the two pictures of the same scene, that is finding for any point on the left image an associated point on the right one. For any such left-right pair of points, this allows, knowing the relative position of the cameras, to evaluate the depth in space of the associated real point. One of the main problem with the implementation of the reconstruction is the presence of obstructed regions, that is significant points or objects that are visible from one optical system but not from the other because of the obstruction of a nearer object. In [5], [11] and [8] the occlusion problem is widely introduced and reviewed.

Stereo correspondence algorithms can be grouped into those producing sparse output, usually feature based, and those giving a dense result. According to the categorization proposed by the important stereo vision taxonomy [13], dense algorithms can be classified as local and global ones. The local ones, or window-based, are lean to speed in the accuracy-speed tradeoff. Global ones are usually slower but more accurate.

An up-to-date various review of stereo vision techniques as well as quick introduction and brief summary to the state-of-the-art, can be found in [7]. In [14] a method for resource-limited systems of stereo vision algorithms. Other important review references in stereo vision methods and for the evaluation of stereo vision algorithms are [3], [9] and [16].

In this work, we present the general theoretical framework for a stereo vision process based on an optimal transport correspondence between the two pictures, therefore it is a global method with a dense result. We consider two cameras looking at the scene having parallel axis and such that the height of any point is the same from the two optical systems. This reduces the problem of building the left-right correspondence to a one-dimensional problem to be carried out on each horizontal section of the pictures, allowing a gain in efficiency. Moreover, the reconstruction becomes totally encoded on a single disparity function.

We consider the two (black and white) pictures as functions \( I_0, I_1 : D \to [0, 1] \) where \( D \subset \mathbb{R}^2 \) is a planar domain and we denote the restrictions to a certain height \( y \) by \( I_0^{(y)}, I_1^{(y)} \). At first we treat the case with no obstructed regions, therefore the two horizontal functions can be seen as finite measures \( \nu_0, \nu_1 \) over \( D^{(y)} = D \cap \mathbb{R} \times \{y\} \) with the same mass, and we can use the toolbox of optimal transport.

After normalization, \( \nu_0, \nu_1 \) are probabilities over \( D^{(y)} \). We use the notation \( \pi^{(0)}, \pi^{(1)} : D^{(y)} \times D^{(y)} \to D^{(y)} \) for the natural projections, and we call admissible (or admissible plan) a probability \( \gamma \in \mathcal{P}(D^{(y)} \times D^{(y)}) \) whose marginals are

\[
\pi^{(0)}_\# \gamma = \nu_0, \quad \pi^{(1)}_\# \gamma = \nu_1.
\]

We denote by \( \Pi(\nu_0, \nu_1) \) the set of admissible plans between \( \nu_0 \) and \( \nu_1 \). The problem of optimal transport, by the Kantorovich’s formulation, consists in minimizing

\[
\langle c, \gamma \rangle := \int c(x, y) d\gamma, \quad (1.1)
\]
with $\gamma$ varying in $\Pi(\nu_0, \nu_1)$, where $c: D(y) \times D(y) \to \mathbb{R}$ is called the cost function. In our work $c$ will be the square difference. It is known (see for example [1]) that under very general assumptions the optimal transport problem has a unique solution and it induces a map from (a subset of the) left domain to the right domain.

Passing to the discrete setting, anyway, the evaluation of the correspondence above is very inefficient. For this reason we consider a regularized version of the optimal transport, we minimize

$$
\langle c, \gamma \rangle - \varepsilon \cdot h(\gamma),
$$

where $\varepsilon$ is a coefficient and $h$ is the entropy function (see §3.2). For any $\varepsilon$ this problem has a unique solution $\gamma^\varepsilon$ in a discrete setting, and for $\varepsilon$ converging to 0 we have $\gamma^\varepsilon \to \gamma^*$, where $\gamma^*$ is the maximal entropy probability among the ones minimizing $\langle c, \gamma \rangle$. The problem of regularized optimal transport has been widely treated, and we use [12, 4, 10] as main references.

Minimizing (1.2) over the space $C$ of admissible plans, is the same of finding the point of $C$ minimizing the Kullback-Leibler divergence from the Gibbs kernel $K^\varepsilon$. Finding the projection $\text{Proj}_{KL}(K^\varepsilon)$ can be done with the so called Sinkhorn’s algorithm that we treat in §4.1, and it has a very good implementation cost.

With respect to our problem, given any probability $\gamma \in \Pi$ we introduce a function $f_\gamma$ (see Definition 4.9) that coincides with the disparity function when $\gamma$ is map induced. In fact $f_\gamma$ works as a good approximation of the disparity function, and our main result is Theorem 5.2 where we prove that the Sinkhorn’s algorithm has an exponential rate of convergence with respect to the functions $f_\gamma$ produced at each step.

We also show that this procedure works very well for artificial non-occluded cartoons (see Figure 6.5), allowing the precise spatial reconstruction with a low implementation cost.

In the last section we also try a first approach to the case where there are occluded regions. This causes the masses of the measures $\nu_0, \nu_1$ to be different, therefore the usual theoretical tools of optimal transport does not work, but we can anyway implement the Sinkhorn's algorithm. We show in Theorem 5.5 and Remark 5.7 that the induced functions $f_\gamma$ converge anyway and the convergence function allows to find the occluded regions in some simple cases. We also build in Figure 6.8 the reconstruction of one of these cases, an artificial cartoon with two object superposing.

2. Preliminaries on stereo vision. We recall the basic geometrical intuition behind stereo vision and introduce the notations that we will use along this work.

2.1. Epipolar geometry. Every camera, and a human eye, can be modelized as a projection on a 2-dimensional plane of a 3-dimensional space. More precisely, in the case of the human eye, the projecting surface is a spherical sector of the retina, but in our discussion we will work with planar projections because they are more fit to treat camera vision.

Stereo vision is the process of matching points between two planar projections of the same space, and the consequent reconstruction of the 3-dimensional scene. The general parametrization of the space through coordinates on two projecting planes is called epipolar geometry (for a wide treatment, see for example [6]).

We call optical centers the two points $O_L, O_R$ that are the centers for the projections of any point $X$ on the two planes. We denote by $e_L$ and $e_R$ the projections of the optical centers $O_R$ and $O_L$ on the left and right plane respectively. For any point $X$ in the space, we denote by $X_L, X_R$ its projections on the two planes, and we call epipolar lines the lines passing through $X_L, e_L$ on the left plane, and $X_R, e_R$ on the right plane. The line $O_LO_R$ through the two optical centers is called optical line.
Figure 2.1. The point $X$ in space is projected on the two planes. The red line on the right is one of the two epipolar lines.

2.2. Transport problem between two images. We denote by $I_0$ and $I_1$ the images from the left and the right optical centers respectively. For both images we have

$$I_j : D \rightarrow [0, 1], \ j = 0, 1,$$

where the images are represented by an intensity function on the planar domain $D \subset \mathbb{R}^2$. Therefore in this approach we are considering images determined by a single intensity function, such as black and white images.

For any fixed $y$ we consider the restricted domain $D(y) := D \cap (\mathbb{R} \times \{y\})$ and the restricted images

$$I_j^y = I_j \mid_{\mathbb{R} \times \{y\}} : D(y) \rightarrow [0, 1], \ j = 0, 1.$$

We consider $I_0^y$ and $I_1^y$ as the densities of two measures over $D(y)$ and we denote by $\nu_0^y, \nu_1^y$, or simply $\nu_0, \nu_1$, the induced measure themselves.

We will at first consider the case where $\nu_0$ and $\nu_1$ have the same total (finite) mass, therefore after normalization they are probability measures over the domain $D(y)$. In §6 we will approach the case where $\nu_0$ and $\nu_1$ have different total mass. We point out that by the nature of the problem the difference in mass is small with respect to the total mass.

If $\nu_0, \nu_1$ are probability measures over $D(y)$, our goal is to find the disparity shift between the two pictures by applying the optimal transport tools, that is by searching the optimal plan that realizes the infimum (3.2). Because of the nature of the problem, we are interested in optimal plans that are induced by transport maps. In fact, the solution to the optimal transport problem is map-induced in this case.

**Proposition 2.1.** Consider a closed interval $J \subset \mathbb{R}$ and two probability measures $\nu_0, \nu_1 \in \mathcal{P}(J)$ that are absolutely continuous with respect to the classic Lebesgue measure over $J$. If we consider the quadratic cost $c : J \times J \rightarrow \mathbb{R}$ such that $c(x, y) = (x - y)^2$, then there exists a unique optimal plan $\gamma$ in $\Pi(\nu_0, \nu_1)$ and it is induced by a map $T : J \rightarrow J$.

**Proof.** If the measures are absolutely continuous with respect to the Lebesgue measure, then they are regular and we conclude by [1, Theorem 1.26].

**Remark 2.2.** We observe that this is precisely our case, where $J = D(y)$ and $\nu_0, \nu_1$ are induced by the densities $I_0^y, I_1^y$. Indeed, $I_0^y, I_1^y$ are continuous almost everywhere and therefore the two induced measures are absolutely continuous with respect to the classic Lebesgue measure.
2.3. **A model of the human optical system.** In our work, we will suppose that the two projection planes are both parallel to the optical line, and this line is considered to be horizontal, in such a way that the vertical coordinate of a same point, is the same in the two projections.

As both image planes and the optical line are parallel, we can consider the distance \( b \) of the image plane (also called projection plane) from the optical system as a single invariant of our system.

For any point \( X \) in space, we denote by \( x_L \) and \( x_R \) the (horizontal) coordinates of its two projections. We are interested in finding the disparity between the position of the object seen from an optical center, and the position from the other one. We suppose that objects “at infinity” are seen without disparity, that is we fix the coordinates on the projecting planes in such a way that parallel lines from the optical centers intercept the projecting plan at the same coordinate \( x_L = x_R \).

![Figure 2.2](image)

**Figure 2.2.** The object \( X \) is projected to \( X_L \) and \( X_R \) respectively from the optical center \( O_L \) and \( O_R \). The coordinate \( x_R \) on the right is set such that \( x_r(X_L') = 0 \), where \( O_RX_L' \) is a segment parallel to \( XO_L \).

As it appear from the image, for an object at distance \( b \) from the optical line, if \( \ell \) is the distance of \( X \) from the optical line and \( \ell_0 \) is the distance \( |O_O| \) between the optical centers, then

\[
x_L - x_R = |X_RX_L'| = \frac{b}{\ell} \cdot \ell_0.
\]  

We remark that there is only one coordinate considered because we are supposing that the optical system is aligned with the vertical ax. This is compatible with the ability of human vision to distinguish ‘up’ from ‘down’. The matching procedure by optimal transport between the left and right projections will be applied at every \( y \) coordinate separately. This speeds up the efficiency of our algorithm, because it scales with \( \lambda^2 \) if \( \lambda \) is the length of the horizontal window that we consider. Instead, if we had considered optimal transport on a 2-dimensional domain, the scaling factor would have been \( \lambda^4 \).

3. **Optimal transport and stereo vision.**

3.1. **Definition of the optimal transport problem.** We recall briefly the formulations of the optimal transport problem. Our main reference is Ambrosio-Gigli guide [1]. We recall that a Polish space is a complete and separable metric space. Consider two Polish spaces \( X, Y \), and a Borel cost function \( c: X \times Y \to \mathbb{R}_{\geq 0} \cup \{+\infty\} \). In the Monge definition of the optimal transport problem, we start from two probability measures \( \nu_0 \) and \( \nu_1 \) in \( X \) and \( Y \) respectively. We are interested in finding (if it exists) a function \( T: X \to Y \) sending \( \nu_0 \) in \( \nu_1 \) and such that the total transport cost is minimized. We recall that the measure \( T_#\nu_0 \) over \( Y \) is defined by

\[
T_#\nu_0(B) = \nu_0(T^{-1}(B)) \quad \forall B \in \mathcal{B}(Y).
\]
Therefore, the optimal transport problem consists in achieving

\[
\inf_{T \# \nu_0 = \nu_1} \int_X c(x, T(x))d\mu,
\]

where the infimum is taken over all the functions \(T\) respecting the constraint \(T \# \nu_0 = \nu_1\).

Kantorovich generalizes this approach, by considering transport plans.

**Definition 3.1.** An admissible transport plan between \(\nu_0\) and \(\nu_1\) is a measure \(\gamma\) over \(X \times Y\) such that \(\pi^X_# \gamma = \nu_0\) and \(\pi^Y_# \gamma = \nu_1\), or equivalently

\[
\gamma(A \times Y) = \nu_0(A) \quad \forall A \in \mathcal{B}(X)
\]

\[
\gamma(X \times B) = \nu_1(B) \quad \forall B \in \mathcal{B}(Y).
\]

We denote the set of admissible transport plans between \(\nu_0\) and \(\nu_1\) by \(\Pi(\nu_0, \nu_1)\).

The optimal transport problem becomes the detection of

\[
\inf_{\gamma \in \Pi(\nu_0, \nu_1)} \int_{X \times Y} c(x, y)d\gamma.
\]

**Definition 3.2.** An optimal plan is an admissible plan achieving the infimum of (3.2).

**Remark 3.3.** We introduce a more compact notation for the coupling on the right. If \(c\) is a cost function and \(\gamma \in \mathcal{P}(X \times Y)\), we define

\[
\langle c, \gamma \rangle := \int_{X \times Y} c(x, y)d\gamma.
\]

In fact, problem (3.2) has a solution under very general conditions.

**Theorem 3.4** ([15, Theorem 4.1]). Consider two measures \(\nu_0, \nu_1\) in \(\mathcal{P}(X)\) and \(\mathcal{P}(Y)\) respectively. If the cost function \(c\) is lower semicontinuous and bounded from below, then there exists an optimal plan \(\gamma\) for the functional \(\gamma \mapsto \langle c, \gamma \rangle\), among all \(\gamma \in \Pi(\nu_0, \nu_1)\).

We are interested in the case where an optimal plan is induced by a transport map \(T\), which means that the optimal plan \(\gamma\) respects \((\text{id} \times T)_\# \nu_0 = \gamma\), or equivalently the solution of (3.1) exists and coincides with the more general solution of (3.2). We saw in Proposition 2.1 and Remark 2.2 that in our cases the opportune conditions are verified for the optimal plan to be induced by a transport map. The same is true in the discrete setting that we approach from now on.

**Remark 3.5.** When we work in a discrete setting, we consider the domain of the probability measures \(\nu_0, \nu_1\) to be \(X = Y = [1, d] \subset \mathbb{N}\). Therefore, \(\nu_0, \nu_1\) are horizontal vectors in \([0, 1]^d\). For the sake of a clearer notation we denote by \(i\) the coordinate on \(X\) and by \(j\) the coordinate on \(Y\).

An admissible transport plan \(\gamma \in \Pi(\nu_0, \nu_1)\) is a \(d \times d\) matrix with positive entries such that

\[
1 \gamma = \mu, \quad \mathbf{1} \gamma^T = \nu,
\]

where \(1 = (1, 1, \ldots, 1)\) is the horizontal vector whose coordinates are all 1. In this presentation, the value \(\gamma_{ij}\) is the mass transported from \(i\) to \(j\) for any \(i, j = 1, \ldots, d\). Furthermore, given a cost function \(c: [1, d]^2 \to \mathbb{R}_{\geq 0}\), the cost coupling becomes

\[
\langle c, \gamma \rangle := \sum_{i,j=1}^{d} c(i, j) \cdot \gamma_{ij}.
\]
3.2. Entropic regularization. We rename the solution of the optimal transport problem

\[ L(\mu, \nu) := \inf_{\gamma \in \Pi(\mu, \nu)} \langle c, \gamma \rangle. \]

By Theorem 3.4 this has a solution if we work with a quadratic cost \( c \). Moreover, the solution is unique in the case of measures induced by densities over a closed interval of \( \mathbb{R} \), as in the case we are considering in treating stereo vision.

For a given admissible plan \( \gamma \), we denote by \( h(\gamma) \) its entropy

\[ h(\gamma) := -\int_{X \times Y} (\log(\gamma) - 1) d\gamma. \]

We consider the optimal transport problem after an entropic regularization, and denote by \( L^\varepsilon \) the regularized infimum,

\[ L^\varepsilon(\mu, \nu) := \inf_{\gamma \in \Pi(\mu, \nu)} \left( \langle c, \gamma \rangle - \varepsilon \cdot h(\gamma) \right). \tag{3.4} \]

For a wide introduction to the entropic regularized problem see [4], [12, §4] and [10], here we recall the main aspects.

**Proposition 3.6.** Given \( \nu_0, \nu_1 \in \mathcal{P}(H) \) for a closed interval \( H \subset \mathbb{R} \), and absolutely continuous with respect to the Lebesgue measure, there exists unique an admissible plan \( \gamma^\varepsilon \) achieving the infimum (3.4), and also

\[ L^\varepsilon(\nu_0, \nu_1) \xrightarrow{\varepsilon \to 0} L(\nu_0, \nu_1). \]

**Remark 3.7.** For a proof of the above result see [4, Theorem 2.7], where is also proved that if \( \gamma \) is the unique optimal plan for \( L(\nu_0, \nu_1) \), then \( \gamma^\varepsilon \) narrowly converges to \( \gamma \).

In the discrete setting \( h(\gamma) = -\sum_{i,j} \gamma_{ij} (\log(\gamma_{ij}) - 1) \) and the result above is also true (see [12, Proposition 4.3]). In particular, if for a given cost function \( c : [1,d]^2 \to \mathbb{R}_{\geq 0} \) we consider the Gibbs kernel

\[ K^\varepsilon \in \mathbb{R}^{d \times d} \text{ s.t. } K^\varepsilon_{ij} := e^{-\frac{c(i,j)}{\varepsilon}} \forall i,j = 1, \ldots, d, \tag{3.5} \]

then the unique solution of (3.4) has the form

\[ \gamma^\varepsilon_{ij} = u^\varepsilon_i K^\varepsilon_{ij} v^\varepsilon_j, \tag{3.6} \]

for some vectors \( u^\varepsilon, v^\varepsilon \in \mathbb{R}^d \). The proof is straightforward after introducing the Lagrangian of (3.4),

\[ \mathcal{L}(\gamma, f, g) := \langle c, \gamma \rangle - \varepsilon h(\gamma) - \langle f, 1^\gamma - \mu \rangle - \langle g, 1_{\gamma^T} - \nu \rangle. \]

**Remark 3.8.** In the discrete setting the optimal plan \( \gamma \) is not necessarily unique. Anyway, by strict concavity of \( h \) there exists unique \( \gamma^* \in \Pi(\nu_0, \nu_1) \) realizing the minimum \( L(\nu_0, \nu_1) \) and maximizing \( h(\gamma^*) \). We have (see [12, Proposition 4.1]) that \( \gamma^\varepsilon \xrightarrow{\varepsilon \to 0} \gamma^* \).

4. Reconstructing disparity shifts with optimal transport. In this section we find an equivalent formulation of the regularized transport problem in terms of a known divergence function. Then, we build an associated disparity function to any transport plan.
4.1. Projecting via the Kullback–Leibler divergence. Divergences are functions that share some properties with distances but are not symmetric nor they respect the triangular inequality. They are frequently used to compare measures and probability distributions.

**Definition 4.1.** If \( P \) is a differentiable manifold, a divergence on \( P \) is a function \( F: P \times P \to \mathbb{R} \) verifying the following properties:

1. \( F(x, y) \geq 0 \) for all \( x, y \in P \);
2. \( F(x, y) = 0 \) if and only if \( x = y \);

One of the most used divergences in probability and statistics is the Kullback–Leibler divergence.

**Definition 4.2.** The Kullback–Leibler divergence, or KL divergence, is defined for any probability measures \( \gamma, \alpha \) over a domain \( Z \). If \( \gamma \) is absolutely continuous with respect to \( \alpha \), then

\[
KL(\gamma|\alpha) := \int_Z \log \left( \frac{d\gamma}{d\alpha} \right) d\gamma,
\]

otherwise \( KL(\gamma|\alpha) = +\infty \).

**Remark 4.3.** In a discrete setting as the one introduced above, \( \gamma \) is always absolutely continuous with respect to \( \alpha \) and the KL divergence becomes

\[
KL(\gamma|\alpha) = \sum_{i,j} \gamma_{ij} \cdot \log \left( \frac{\gamma_{ij}}{\alpha_{ij}} \right).
\]

The KL divergence behaves similarly to the regularized cost if we consider the Gibbs kernel \( K^\varepsilon \) introduced in (3.5). Indeed,

\[
\varepsilon \cdot (KL(\gamma|K^\varepsilon) - 1) = \langle c, \gamma \rangle - \varepsilon \cdot h(\gamma),
\]

therefore minimizing the right side of the above equation for \( \gamma \in \Pi(\nu_0, \nu_1) \), is equivalent to minimize the KL divergence with \( K^\varepsilon \) for \( \gamma \) in the same space.

**Remark 4.4.** Observe that \( K^\varepsilon \) is not a probability distribution. It is however possible to define the KL divergence, although it does not have all the divergence properties above listed. Therefore, the optimal transport regularized problem (3.4) becomes the problem of finding the projection with respect to KL of the Gibbs kernel \( K^\varepsilon \) on \( \Pi(\mu, \nu) \).

We introduce two affine subspaces of the probability space \( P([1, d]) \) where the conditions (3.3) are verified.

**Definition 4.5.** For any \( \mu, \rho \in P([1, d]) \) we denote by \( C^1_\mu \) the subspace of \( P([1, d]^2) \) of probability measures \( \gamma \) verifying \( 1\gamma = \mu \). Analogously, we denote by \( C^2_\nu \) the subspace of \( \gamma \) verifying \( 1\gamma^T = \rho \).

**Remark 4.6.** Observe that \( \Pi(\nu_0, \nu_1) = C^1_{\nu_0} \cap C^2_{\nu_1} \). In order to state an algorithmic implementation for the resolution of the regularized optimal transport, we consider alternatively the projections on \( C^1_{\nu_0} \) and \( C^2_{\nu_1} \):

\[
\gamma^{(2k+1)} := \text{Proj}_{C^1_{\nu_0}}(\gamma^{(2k)}),
\]

\[
\gamma^{(2k+2)} := \text{Proj}_{C^2_{\nu_1}}(\gamma^{(2k+1)}),
\]

for any \( k \in \mathbb{N} \), after initializing the procedure with \( \gamma^{(0)} := K^\varepsilon \).

**Proposition 4.7.** For any function \( K \) over \([1, d]^2\) and any probability \( \mu \in P([1, d]) \), the plan \( \pi \in C^1_\mu \) minimizing \( KL(\pi|K) \) is

\[
\pi_{ij} := \overline{u}_i K_{ij},
\]

where the vector \( \overline{u} \in \mathbb{R}^d \) is the unique vector such that \( \overline{u} \) respects \( 1\overline{u} = \mu \).
Analogously the same is true for $C^2_\rho \ni \gamma := (K_j \pi_j)$ where $\pi$ is the unique vector such that $\gamma$ respects $\mathbf{1}\gamma^T = \rho$.

Proof. We prove the first result. The Lagrangian of $\inf_{\gamma \in C^2_\rho} \text{KL}(\gamma|K)$ is

$$L(\gamma, f) = \sum_{i,j}(\gamma_{ij} \log(\gamma_{ij}) - \gamma_{ij} \log(K_{ij})) - \langle \beta, \mathbf{1}\gamma - \mu \rangle,$$

where $\beta \in \mathbb{R}^d$ is a variable $d$-dimensional vector. If we derive in $\gamma_{ij}$ we get

$$\frac{\partial L}{\partial \gamma_{ij}} = \log(\gamma_{ij}) + 1 - \log(K_{ij}) - \beta_i.$$ 

If we impose $\frac{\partial L}{\partial \gamma_{ij}} = 0$ for any $i,j \in [1,d]$, then we must have

$$\gamma_{ij} = K_{ij} \cdot e^{\beta_i - 1}.$$ 

We define $\pi_i := e^{\beta_i - 1}$ and we get the matrix form in the theorem. The condition $\mathbf{1}\gamma = \mu$ is necessary, but by construction is now also sufficient and gives the unicity of the solution. \qed

This motivates the well known Sinkhorn’s algorithm. After initializing the procedure with the vectors $u^{(0)} = v^{(0)} = \mathbf{1}_d$, we have

$$u^{(k+1)} := \nu_0 \odot (K^\varepsilon \cdot u^{(k)}),$$

$$v^{(k+1)} := \nu_1 \odot ((K^\varepsilon)^T \cdot u^{(k+1)}),$$

where the division on the right sides is the entry-wise division. And with this notation we have

$$\left(\gamma^{(2k+1)}_{ij}\right) = \left(u^{(k+1)}_{i} K^\varepsilon_{ij} v^{(k)}_{j}\right)$$

and

$$\left(\gamma^{(2k)}_{ij}\right) = \left(u^{(k)}_{i} K^\varepsilon_{ij} v^{(k)}_{j}\right).$$

As $C^{1}_{\nu_0}$ and $C^{2}_{\nu_1}$ are affines, by [2] we know that the iterative process converges to $\text{Proj}^{\text{KL}}_{\mathbb{H}(\nu_0, \nu_1)}(K^\varepsilon)$ which is precisely the matrix $\gamma^\varepsilon$ as expressed in (3.6).

Remark 4.8. Observe that the vectors $\nu^\varepsilon, v^\varepsilon$ depend on the initializing vectors $u^{(0)}, v^{(0)}$ but the unique optimal plan $\gamma^\varepsilon$ does not.

4.2. Disparity functions. Any stereo vision reconstruction of 3-dimensional scene is equivalent to the datum of a disparity shift at any point of the right projection $I_1$, that is a vector encoding the shift of that point when passing from the right optical system to the left one. In our setting, the two optical systems are vertically aligned, therefore the disparity shift is a scalar positive value giving the horizontal shift (to the right).

Therefore, given a spatial configuration $C$ of some objects and the associated stereo vision picture pair $I_0, I_1$, at any $y$ coordinate the spatial reconstruction must be a disparity function $f : S \rightarrow \mathbb{R}_{\geq 0}$, where $J$ is a closed interval in $\mathbb{R}$ and $S \subset J$.

In this section we work in the discrete setting, therefore the domain of the disparity functions is a subset of $J = [1,d]$. Furthermore, we consider disparity functions resulting from images where there is no occlusion, meaning that there are no areas of the depicted objects that are invisible from one of the optical systems because of the superposition of other objects. We will consider cases with occlusions in §6.

We associate to any plan $\gamma \in \mathcal{P}([1,d]^2)$ a function $f_\gamma$ and we will prove below (see Proposition 4.15) that it is a non-negative disparity function in the case of stereo vision induced plans.

Definition 4.9. If $\gamma \in \mathcal{P}([1,d]^2)$, for any $i \in [1,d]$ such that $\mathbf{1}\gamma(i) > 0$, we define the following function,

$$f_\gamma(i) := \frac{\sum_j \gamma_{ij} \cdot j}{\mathbf{1}\gamma(i)} - i.$$
Definition 4.10. We call object any compact closed submanifold \( X \) in \( \mathbb{R}^3 \) whose boundary \( \partial X \) is smooth \( \mathcal{L}_{\partial X} \)-a.e. and \( \mathcal{L}_{\partial X} \) is the canonical Lebesgue measure on \( \partial X \).

Because of the aligned disposition of the optical systems, we consider any horizontal section of the configuration, and its left and right projection to the optical line.

Consider a stereo vision picture pair \( I_0, I_1 : [1, d] \to [0, 1] \) obtained in the discrete setting from a configuration \( C \). If \( I_0, I_1 \) have the same finite mass at any \( y \), we can normalize, obtaining \( i_0(y), i_1(y) \in \mathcal{P}([1, d]) \). In order to apply the optimal transport theorems, we consider the cost function \( c : [1, d]^2 \to \mathbb{R}_{\geq 0} \) defined by

\[
c(i, j) := (i - j)^2.
\]

Lemma 4.1. Consider \( \gamma \) an optimal plan for the problem \( L(\nu_0, \nu_1) = \inf \langle c, \gamma \rangle \), where the infimum is taken in \( \Pi(\nu_0, \nu_1) \), and \( \gamma_{ij} > 0 \) and \( \gamma_{i'j'} > 0 \) for some \( i, i', j, j' \in [1, d] \). If \( i' > i \), then \( j' \geq j \).

Equivalently, if \( \gamma_{ij} > 0 \) then there are no \( i' > i \) and \( j' < j \) such that \( \gamma_{i'j'} > 0 \).

Proof. Let’s suppose that there exists \( i' > i \) and \( j' < j \) such that \( \gamma_{ij} > 0 \) and \( \gamma_{i'j'} > 0 \). We want to show that the plan \( \gamma \) is not optimal, contradicting the hypothesis. In order to do this, we introduce

\[
\gamma_{\text{min}} := \min(\gamma_{ij}, \gamma_{i'j'}).\]

and define another plan \( \gamma' \in \Pi(\nu_0, \nu_1) \), equal to \( \gamma \) except that

\[
\gamma'_{ij} := \gamma_{ij} - \gamma_{\text{min}} \\
\gamma'_{i'j'} := \gamma_{i'j'} + \gamma_{\text{min}} \\
\gamma'_{i'j'} := \gamma_{i'j'} - \gamma_{\text{min}} \\
\gamma'_{ij} := \gamma_{ij} + \gamma_{\text{min}}
\]

In order to show that \( \langle c, \gamma' \rangle < \langle c, \gamma \rangle \) it suffices to observe that

\[
(i' - j')^2 + (i - j)^2 > (i' - j)^2 + (i - j')^2.
\]

\[\square\]

Definition 4.11. We say that a point \( P \) of an object \( X \) of the configuration projects to the left (or the right) optical system if the segment \( O_L P \) (or \( O_R P \)) does not intersect any object except at \( P \).

In this case, we say that the coordinate \( x_L(P) \) (or \( x_R(P) \)) represents the projected point \( P \).

Definition 4.12. We call occluded regions those subsets of the space that are ‘visible’ from one of the two optical systems but not from the other. More precisely

\[
H_0 := \{ x \in [1, d] \mid x = x_R(P) P \text{ is not represented on the left o.s.} \};
\]
\[
H_1 := \{ x \in [1, d] \mid x = x_L(P) P \text{ is not represented on the right o.s.} \}.
\]

In the example below, we see that areas \( H_j \) correspond to elements that are covered by a closer (to the optical system) object, and therefore are invisible for one eye or the other. For an introduction to the occlusion problem see for example [2, 3].

We want to define the disparity function associated with a spatial configuration of a finite number of objects. We start by considering very simple configurations.

Definition 4.13. We call non-occluded cartoon a configuration where any object is included in a plane or a line parallel to the optical line, and every point on any object is represented on both optical systems.
We will drop the ‘non-occluded’ hypothesis in the following section, adapting our methods to cases with non-empty occluded regions.

Given a non-occluded cartoon $C$, we consider the horizontal section at a chosen height $y$ and define an associated disparity function $f^{(C,y)}$ that we denote simply $f^{(C)}$ when there is no risk of confusion. We consider the subset $S \subset [1,d]$ of $x_R$ representing a point $P$ on any object of $C$.

**Definition 4.14.** If $x_L$ and $x_R$ are defined as in §2.3, for any $x_R \in S$ representing a point $P$, we have

$$f^{(C)}(x_R) := x_L(P) - x_R(P).$$

By definition, for any non-occluded cartoon $C$ and any $y$ coordinate, $f^{(C,y)}$ is a non-negative function.

Consider a stereo vision picture pair $I_0, I_1 : [1,d] \to [0,1]$ obtained in the discrete setting from a non-occluded cartoon $C$. By construction, $I_0, I_1$ have the same finite mass at any $y$, and we denote again by $\nu_0, \nu_1 \in \mathcal{P}([1,d])$ the probabilities obtained after the normalization.

**Proposition 4.15.** Given a non-occluded cartoon $C$ and any optimal plan $\gamma$ between the induced probabilities $\nu_0, \nu_1$ for the cost function $c$ defined above, we have $f_\gamma = f^{(C)}$ and therefore $f_\gamma$ is a non-negative function.

**Proof.** We prove inductively that for any point $i \in S = \{x \in [1,d] | \nu_0(x) > 0\}$, if $i$ represents a point $P$ of the configuration, that is $i = x_R(P)$, then $\gamma_{ij} = \nu_0(i)$ if $j = x_L(P)$, and $\gamma_{ij} = 0$ for any other $j$.

This is true for the minimum $i \in [1,d]$ such that $\nu_0(i) > 0$, because if $i = x_R(P)$ then $j = x_L(P)$ is the minimum $j$ such that $\nu_1(j) > 0$ and $\nu_1(j) = \nu_0(i)$ by construction. And we conclude with Lemma 4.1.

Suppose now that the proposition is true up to $i_0$, and $i_0 = x_R(P)$, $j_0 = x_L(P)$. Consider $i_1 := \text{argmin}\{i > i_0, \nu_0(i) > 0\}$, and $j_1 := \text{argmin}\{j > j_0, \nu_1(j) > 0\}$. By the absence of obstruction we must have $i_1 = x_R(Q)$ for some point $Q$ of the configuration such that $j_1 = x_L(Q)$. Moreover, by Lemma 4.1 the mass in $j_1$ must be ‘filled’ with the mass ‘coming from’ $i_1$, therefore $\gamma_{i_1,j_1} = \nu_0(i_1) = \nu_1(j_1)$. □

**5. Convergence results.** We show the convergence of the estimates for the optimal plan and also of the associated disparity functions. Furthermore, we introduce here a slight variation of the problem we considered above by that the target measure $\nu_1$ is a probability measure while the total
mass of the source measure is $m(\nu_0) > 1$. This will in particular account for the difference among the two pictures due to the occlusion of some objects by other objects that are nearer to the optical system. The case where the target measure $\nu_1$ carries the greater mass is theoretically equivalent to the first case, but with slight differences in the applied consequences in stereo vision settings.

5.1. Convergence of associated functions. In order to treat the convergence of Sinkhorn’s algorithm, we introduce for any vector $w \in \mathbb{R}^d$ the variation norm

$$\|w\|_{\text{var}} := \max_i w_i - \min_i w_i,$$

and for any two vectors $u, u' \in \mathbb{R}_{>0}^d$, the Hilbert metric

$$d_H(u, u') := \|\log(u) - \log(u')\|_{\text{var}}.$$  

Observe that in order for $d_H$ to be a metric, we have to quotient out the relation $u \sim u'$ if $u = \lambda \cdot u'$ for some $\lambda \in \mathbb{R}_{>0}$. Over $\mathbb{R}_{>0}^d/\sim$, $d_H$ respects the triangle inequality and it is a distance.

For every positive matrix $K \in \mathbb{R}_{>0}^{d \times d}$, we define

$$\eta(K) := \max_{i,j,i',j'} \frac{K_{ij}}{K_{i'j'}} \cdot \frac{K_{i'j'}}{K_{ij}}$$

and moreover

$$\lambda(K) := \frac{\sqrt{\eta(K)} - 1}{\sqrt{\eta(K)} + 1}.$$  

Therefore $0 < \lambda(K) < 1$ and by the following theorem $K$ acts as a contraction on the cone of positive vectors.

**Theorem 5.1.** For $K$ as above and $u, u' \in \mathbb{R}_0^d$,

$$d_H(Ku, Ku') \leq \lambda(K) \cdot d_H(u, u').$$

For a proof see [12, Theorem 4.1]. Moreover, as a direct consequence of the theorem above we can state a convergence rate result for the sequences $u^{(k)}$ and $v^{(k)}$ converging to $u^\varepsilon$, $v^\varepsilon$ that we treated in §4.1.

We observe that if $u^\varepsilon(i) = 0$ for some $i \in [1, d]$, then $\nu_0(i) = 0$ and $u^{(k)}(i) = 0$ for any $k \geq 1$. And the same is true if $v^\varepsilon(j) = 0$ for $\nu_1(j) = 0$ and $v^{(k)}(j) = 0$ if $k \geq 1$. Therefore the convergence on null coordinates is completed at the first iteration, and for the evaluation of the distance between two vectors we can focus only on non-null coordinates. Thus, up to considering a coordinate subspace $\mathbb{R}^d \subset \mathbb{R}^d$ we can suppose that $d_H(u^{(k)}, u^\varepsilon)$ and $d_H(v^{(k)}, v^\varepsilon)$ are well defined for any $k \in \mathbb{N}_{>0}$, or equivalently that their coordinates are everywhere non-null.

**Lemma 5.1.** If $u^{(k)}, v^{(k)}$ are the convergent series of vectors induced by the Sinkhorn’s algorithm as explained above, we have the following convergence rate

$$d_H(u^{(k)}, u^\varepsilon) = O(\lambda(K^\varepsilon)^{2k})$$

$$d_H(v^{(k)}, v^\varepsilon) = O(\lambda(K^\varepsilon)^{2k}).$$

**Proof.** For any triple of everywhere non-null vectors $u, u', a \in \mathbb{R}_{>0}^d$, we have that $d_H(u, u') = d_H(a \odot u, a \odot u')$. Moreover, because of the iteration rules of the Sinkhorn’s algorithm, we know that

$$d_H(u^{(k)}, u^\varepsilon) = d_H(\nu_0 \odot (K^\varepsilon v^{(k-1)}), \nu_0 \odot (K^\varepsilon v^\varepsilon))$$

$$= d_H(K^\varepsilon v^{(k-1)}, K^\varepsilon v^\varepsilon)$$

$$\leq \lambda(K) \cdot d_H(v^{(k-1)}, v^\varepsilon).$$
In the same way we can prove \( d_H(v^{(k-1)}, v^\varepsilon) \leq \lambda(K^\varepsilon) \cdot d_H(u^{(k-1)}, u^\varepsilon) \), and this concludes the proof. \( \square \)

We are able to translate the convergence rate result above in a result about the convergence of disparity functions. We already considered the unique admissible plan \( \gamma^\varepsilon \) minimizing \( L_\varepsilon(\nu_0, \nu_1) \) and the series of plans \( \gamma^{(2k+1)} = u^{(k+1)}K^\varepsilon v^{(k)} \) for any \( k \in \mathbb{N}_{>0} \), obtained via the Sinkhorn’s algorithm. We define

\[
\begin{align*}
    f_\varepsilon &:= f_{\gamma^\varepsilon} \\
    f_{(k)} &:= f_{\gamma^{(2k+1)}}
\end{align*}
\]

We show that \( f_{(k)} \) converges to \( f_\varepsilon \) and estimate the convergence rate.

**Theorem 5.2.** For \( \varepsilon \to 0 \), we have \( f_{(k)} \to f_\varepsilon \) and moreover

\[
\|f_{(k)} - f_\varepsilon\|_\infty = O(\lambda(K^\varepsilon)^{2k}).
\]

**Proof.** Observe that for any \( i \in [1, d] \),

\[
u_0(i) = \sum_j K^\varepsilon_{ij} v^\varepsilon_j
\]

for any \( k \in \mathbb{N} \). Therefore we can rewrite the disparity function for \( \gamma^{(2k+1)} \) evaluated at any \( i \) gives

\[
    f_{(k)}(i) + i = \sum_j u^{(k+1)} K^\varepsilon_{ij} v^{(k)}_j \cdot j = \sum_j K^\varepsilon_{ij} v^{(k)}_j \cdot j
\]

because \( \nu_0(i) = 1 \gamma^{(2k+1)}(i) \). By Lemma 5.1 if we call \( d_H(v^{(k)}, v^\varepsilon) =: A \), then

\[
e^{-A} \cdot \sum_j K^\varepsilon_{ij} v^\varepsilon_j \cdot j \leq \sum_j K^\varepsilon_{ij} v^{(k)}_j \cdot j \leq e^A \cdot \sum_j K^\varepsilon_{ij} v^\varepsilon_j \cdot j.
\]

By straightforward calculations this implies

\[
|f_{(k)}(i) - f_\varepsilon(i)| \leq d \cdot (1 - e^{-A}),
\]

and again by Lemma 5.1 this implies the theorem. \( \square \)

### 5.2. Shifted projections.

In the case where \( \nu_1 \) is a probability measure but \( \nu_0 \) is not, and it has a total mass \( m_0 := m(\nu_0) > 1 \), the main problem in evaluating the regularized optimal plan is that \( \Pi(\nu_0, \nu_1) = C^1_{\nu_0} \cap C^2_{\nu_1} = \emptyset \). In the following we are going to use the Sinkhorn’s procedure of successive projections, and show that we will obtain two separate convergent series.

Observe that the measure \( \tilde{\nu}_0 := \frac{1}{m_0} \cdot \nu_0 \) is a probability measure by definition. Moreover, observe that any element of \( C^1_{\nu_0} \) has mass \( m_0 \) and any element of \( C^1_{\tilde{\nu}_0} \) is a probability measure.

**Definition 5.3.** We call shifting \( \text{sh} : C^1_{\nu_0} \to C^1_{\tilde{\nu}_0} \) the map that sends \( \mu \mapsto \frac{1}{m_0} \cdot \mu \). When there is no risk of confusion we use the notation \( \tilde{\mu} = \text{sh}(\mu) \).

**Remark 5.4.** Observe that this map preserves the affine structure with respect to the vector space of 0 mass functions. Moreover, \( \text{sh} \) is invertible as a map of affine spaces.

We work again in the discrete setting where \( \nu_0, \nu_1 \) are defined over \([1, d]\). For any measure \( \alpha \) over \([1, d]^2\) consider the two sequences

\[
\begin{align*}
    \gamma^{(2k+1)} &= \text{Pro}_{C^1_{\nu_0}} K_{\alpha}^{\nu_0}(\gamma^{(2k)}) \\
    \gamma^{(2k+2)} &= \text{Pro}_{C^1_{\nu_1}} K_{\alpha}^{\nu_1}(\gamma^{(2k+1)})
\end{align*}
\]

12
for any \( k \in \mathbb{N} \), after the initialization \( \gamma_{\alpha}^{(0)} := \alpha \), analogous to what we did in Remark 4.6 for the Gibbs kernel \( K^\circ \). In this case we obtain two separate convergence results.

**THEOREM 5.5.** In the case above, we have that the even sequence converges

\[
\gamma_{\alpha}^{(2k)} \xrightarrow{k \to +\infty} \left( \text{Proj}_{\Pi(\nu_0, \nu_1)}^{\text{KL}}(\alpha) \right) \in C^2_{\nu_1},
\]

and the odd sequence converges too, but to another limit

\[
\gamma_{\alpha}^{(2k+1)} \xrightarrow{k \to +\infty} \text{sh}^{-1} \left( \text{Proj}_{\Pi(\nu_0, \nu_1)}^{\text{KL}}(\alpha) \right) \in C^1_{\nu_0}.
\]

**Proof.** We are going to show that for any measure \( \mu \in C^1_{\nu_0} \), minimizing the KL divergence \( \text{KL}(\mu|\alpha) \) is equivalent to minimizing \( \text{KL}(\bar{\mu}|\alpha) \). Indeed,

\[
\text{KL}(\mu|\alpha) = \int \log \left( \frac{d\mu}{d\alpha} \right) d\mu = \int \log \left( \frac{m_0 \cdot d\bar{\mu}}{d\alpha} \right) m_0 \cdot d\bar{\mu} = \int m_0 \cdot \left( \log(m_0) + \log \left( \frac{d\bar{\mu}}{d\alpha} \right) \right) d\bar{\mu} = m_0 \cdot \left( \log(m_0) + \text{KL}(\bar{\mu}|\alpha) \right).
\]

Moreover, \( \bar{\mu} \) lies in \( C^1_{\nu_0} \) which intersects with \( C^2_{\nu_1} \) at \( \Pi(\nu_0, \nu_1) \). Therefore the sequence of probabilities \( \gamma_{\alpha}^{(1)} , \gamma_{\alpha}^{(2)} , \gamma_{\alpha}^{(3)} , \gamma_{\alpha}^{(4)} , \ldots \) respects the conditions of the sequence in Remark 4.6 with initialization equal to \( \alpha \). Therefore both \( (\gamma_{\alpha}^{(2k+1)})_k \) and \( (\gamma_{\alpha}^{(2k)})_k \) converge to \( \text{Proj}_{\Pi(\nu_0, \nu_1)}^{\text{KL}}(\alpha) \), proving the theorem. \( \Box \)

**REMARK 5.6.** Observe that the limit of \( \gamma_{\alpha}^{(2k+1)} \) lies in the subset of \( C^1_{\nu_0} \) of minimal KL divergence from \( C^2_{\nu_1} \), that is \( \text{sh}^{-1}(\Pi(\nu_0, \nu_1)) \).

**REMARK 5.7.** If \( \gamma \in \Pi(\nu_0, \nu_1) \) and \( \bar{\gamma} = \text{sh} \gamma \), then the associated functions \( f_\gamma \) and \( f_{\bar{\gamma}} \) coincide. Therefore in the succession above, also the disparity function converges, and the converge rate result of Theorem 5.2 is still true.

6. **Numerical results.** First we consider the case of a pair of probability measures obtained by taking two artificially drawn images that are shifted one respect to the other and have no obstructions. Then we also apply the same method for cartoons with a simple obstruction, that is one object covers another one on the same scene.

6.1. **An example of non-occluded cartoon.** In this section we consider the discrete domain \( D = [1, d] \times [1, h] \subset \mathbb{N}^2 \). Therefore the two images are functions

\[
I_j : [1, d] \times [1, h] \to [0, 1], \quad j = 0, 1,
\]

and for every \( y \in [1, h] \) coordinate, \( f_j^{(y)} : [1, d] \to [0, 1] \) are the intensity functions that determine (after normalization) the probabilities \( \nu_0^{(y)} , \nu_1^{(y)} \). We consider the case of a non-occluded cartoon as by Definition 4.13. This represents a first modelization of a real configuration of objects where no-object is superposed to any other from both of the optical systems.

By Proposition 4.15 in the case of a quadratic cost and a non-occluded cartoon, the only optimal plan \( \gamma \) is the one induced by the natural matching between the two stereo vision pictures.

In order to evaluate this matching plan for every horizontal line, our procedure consists in implementing the Sinkhorn’s algorithm described in §4.1 to every height \( y \), because the regularized
transport solution is a good approximation of $\gamma$ if the regularization coefficient is sufficiently small (see Remark 3.8). Other than the $I_j$ functions, we have to introduce the entropic coefficient $\varepsilon$ and the number of iterations niter.

Algorithm 1 Computation of the transport plan between $I_0$ and $I_1$ via the Sinkhorn’s method

**Input** $I_0$, $I_1$, $\varepsilon$, niter.

Set $d = \text{size}(I_0, 2) = \text{size}(I_1, 2)$, and $h = \text{size}(I_0, 1) = \text{size}(I_0, 1)$;
Define the cost matrix $c \in \mathbb{R}^{d \times d}$ such that $c(i, j) := (i - j)^2$;
Define the kernel $K^{\varepsilon}(i, j) := e^{-\frac{c(i, j)}{\varepsilon}}$;
Initialize $v^{(0)} := \text{ones}(1, d)$;

for $y = 1, \ldots, h$ do
Set $\nu_0 := I_0(y, :)$ and $\nu_1 := I_1(y, :)$;
for $k = 1, \ldots, \text{niter}$ do
$u^{(k)} := \nu_0 \oslash (K^{\varepsilon}v^{(k-1)})$;
$v^{(k)} := \nu_1 \oslash ((K^{\varepsilon})^T u^{(k)})$;
end for
Set $u^{\varepsilon} := u^{(\text{niter})}$ and $v^{\varepsilon} := v^{(\text{niter})}$;
Set $\gamma^{\varepsilon}(:, :, y) = \text{diag}(u^{\varepsilon})K^{\varepsilon} \text{diag}(v^{\varepsilon})$.
end for

We develop this with the cartoon in the image below, taking three horizontal sections as examples.

**Figure 6.1.** Two pictures of the same non-obstructed cartoon taken from two different optical systems. We marked three horizontals sections for $y = 40, 60, 80$, the rectangular object $X$ and we also marked in blue the intersection $D^{(80)} \cap X$.

**Figure 6.2.** The intensity functions $I_0^{(60)}$ (in blue) and $I_1^{(60)}$ (in orange), at the horizontal section $y = 60$.

Implementing our algorithm with niter $= 10^5$ and $\varepsilon = 0.1$ we obtain the optimal transport plans and the disparity functions depicted in Figure 6.3. The disparity results agree with the cartoon configuration we used. Finally we develop a 3-dimensional reconstruction of the cartoon, that only depends
on the disparity functions and the geometrical invariants of the optical system.

Figure 6.3. An approximation of the optimal transport plan matrices for the marked lines $y = 80, 60, 40$, after a sigmoid passage that emphasizes the non-null entries. The marked segment on the first image is the intersection $X \cap D(80)$. The matrices are obtained with the log version of the Sinkhorn's algorithm described below.

Figure 6.4. The disparity functions obtained at $y = 80, 60, 40$. See Definition 4.9.

6.2. From shifted projections to the recovery of occluded regions. Starting from the results of §5.2, it is possible to determine the occluded regions of a given configuration if they exist. We consider in particular the case of a cartoon configuration where $H_0 = \emptyset$ while $H_1 \neq \emptyset$, meaning that there is some point represented in the left optical system that is not represented in the right optical system, but the inverse is not true. If $H_1 \neq \emptyset$ then $m(I_0^{(y)}) \geq m(I_1^{(y)})$ for some $y$.

Differently from the non-obstructed case, we obtain that the two measures $\nu_0^{(y)}$ and $\nu_1^{(y)}$ may have different masses. Therefore, after dividing by $m(I_0^{(y)})$ $\nu_1$ is a probability measure while $m(\nu_0) \geq 1$. We focus on those $y$ where the last inequality is strict, and therefore the mass quotient $\varphi^{(y)} := m(I_1^{(y)})/m(I_0^{(y)})$ is strictly lower than 1. When there is no risk of confusion we denote this quotient by $\varphi$.

Definition 6.1. We call compression coefficient of a point $x \in [1, d]$ de derivative of $f_\gamma$ if it is well defined in $x$.

We observe that by the results of §4.2 if $\nu_0, \nu_1$ are two probabilities then the compression coefficient is 0 everywhere it is defined. If $m(\nu_0) > 1$ then, as stated in Remark 6.7, the disparity function $f_\gamma$ can be obtained by considering the optimal transport problem between $\tilde{\nu}_0$ and $\nu_1$. This means that the compression coefficient can take non-null values.

We consider the leftmost object in our configuration and denote it by $X$. As we are matching $\tilde{\nu}_0$ instead of $\nu_0$, then the mass of $X$ is reduced by a $\varphi$ factor. Using Lemma 4.1 we observe that $\tilde{\nu}_0|_X$ ‘fills’ the corresponding area of $\nu_1$ slower than in the non-obstructed case. We state without proof that
Figure 6.5. A three dimensional reconstruction after the estimate (line by line) of the disparity shift. We supposed that $\ell_0 = 10$ and $b = 1000$ (see Figure 2.2), and we have a conversion coefficient $\beta = 2$ between pixels and the length unit, therefore $\ell = \frac{b \cdot \ell_0}{\beta (x_L - x_R)}$.

The new compression coefficient at points in $X$ approximates $\phi - 1$.

If $i_0$ is the coordinate of the leftmost point of $X$, then $f_\gamma(i_0)$ recovers correctly the shift of $X$. This means that we can recover also the region occluded by $X$. Indeed, if we know the rightmost point $i_1$ of $X$, then by checking how much of the mass of $I_1$ lies at the right of $i_1 + f_\gamma(i_0)$ we can recover the area that is covered by $X$. This means recovering the point $i_2$ such that the segment $[i_1 + 1, i_2]$ is occluded by $X$. We consider then $\nu'_0 := \nu_0 - \nu_0$ $[i_1 + 1, i_2]$. If $m(\nu'_0) = 1$ then we have recovered all the occluded area; if not, we repeat the same procedure by eliminating the object $X$ both from $\nu_0$ and $\nu_1$.

In the example below, we considered a case with $H_0 = \emptyset$ and $H_1 \neq \emptyset$, in particular focusing at $y = 30$ we obtain an occluded region and recover it.

Figure 6.6. In this figure we added a fourth object that has a superposition with the circle. On the right we show the transport plan at $y = 30$. 

16
Figure 6.7. The function obtained from the optimal plan at $y = 30$. As we can see the disparity shift for the leftmost object is $9$.

Figure 6.8. We show the disparity function, where different colors correspond to different disparity values, and the dark blue to points with no mass, therefore including also occluded regions. On the right we have the three dimensional reconstruction.

Remark 6.2. We observe from the image below that the value of the compression coefficient is not constant over the $X$ points of $[1, d]$. This is a consequence of the discretization of the domain. If we want to evaluate $f_{\gamma}$ at $X$ we start by imposing, without loss of generalities, that $i_0 = 1$. Then, we have two cases for any point $i$ in $X$.

If $\lfloor \varphi (i - 1) \rfloor < \lfloor \varphi i \rfloor$, then

$$f_{\gamma}(i) = \frac{1}{\varphi} ((\varphi i) \cdot \lfloor \varphi i \rfloor + (1 - \{ \varphi \cdot (i - 1) \}) \cdot \lfloor \varphi i \rfloor) - i$$

$$= \left(1 - \frac{1}{\varphi}\right) \cdot \lfloor \varphi i \rfloor.$$

While, if $\lfloor \varphi (i - 1) \rfloor = \lfloor \varphi i \rfloor$, then

$$f_{\gamma}(i) = \lfloor \varphi i \rfloor - i.$$

The derivative operator in a discrete setting is defined as $\Delta f(i) := f(i + 1) - f(i)$. We observe that therefore $\Delta f_{\gamma}$ takes the same value for two successive entries $i, i + 1$ only when this value is $1 - \frac{1}{\varphi}$ and this allows to recover the $\varphi$ value and to check our implementation.
Figure 6.9. The graph of the discrete derivative $\Delta f_\gamma$. We recover from this that $1 - \frac{1}{\varphi} \approx -0.05637$ and therefore $\varphi \approx 0.95$, which coincides with the true estimate.

REFERENCES

[1] Luigi Ambrosio and Nicola Gigli. A User's Guide to Optimal Transport, pages 1–155. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[2] Lev M Bregman. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. USSR computational mathematics and mathematical physics, 7(3):200–217, 1967.

[3] Myron Z Brown, Darius Burschka, and Gregory D Hager. Advances in computational stereo. IEEE transactions on pattern analysis and machine intelligence, 25(8):993–1008, 2003.

[4] Guillaume Carlier, Vincent Duval, Gabriel Peyré, and Bernhard Schmitzer. Convergence of entropic schemes for optimal transport and gradient flows. SIAM Journal on Mathematical Analysis, 49(2):1385–1418, 2017.

[5] Davi Geiger, Bruce Ladendorf, and Alan Yuille. Occlusions and binocular stereo. International Journal of Computer Vision, 14(3):211–226, 1995.

[6] Richard Hartley and Andrew Zisserman. Multiple view geometry in computer vision. Cambridge university press, 2003.

[7] Shangmeng He, Ziyuan Tong, Guobin Ma, Mingyun Fan, Ling Lingzhou, and Shoufeng Tang. Research on stereo vision matching algorithm for rescue robot. In 2017 International Conference on Robotics and Automation Sciences (ICRAS), pages 35–38. IEEE, 2017.

[8] Shafik Huq, Andreas Koschan, and Mongi Abidi. Occlusion filling in stereo: Theory and experiments. Computer Vision and Image Understanding, 117(6):688–704, 2013.

[9] Nalpantidis Lazaros, Georgios Christou Sirakoulis, and Antonios Gasteratos. Review of stereo vision algorithms: from software to hardware. International Journal of Optomechatronics, 2(4):435–462, 2008.

[10] Christian Léonard. From the schrödinger problem to the monge–kantorovich problem. Journal of Functional Analysis, 262(4):1879–1920, 2012.

[11] Xiaoyong Lin, Yu Liu, and Wenzhan Dai. Study of occlusions problem in stereo vision. In 2008 7th World Congress on Intelligent Control and Automation, pages 5062–5067. IEEE, 2008.

[12] G. Peyré and M. Cuturi. Computational Optimal Transport: With Applications to Data Science. Foundations and trends in machine learning. Now, the essence of knowledge., 2019.

[13] Daniel Scharstein and Richard Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. International journal of computer vision, 47(1):7–42, 2002.

[14] Beau Tippetts, Dah Jye Lee, Kirt Lillywhite, and James Archibald. Efficient stereo vision algorithms for resource-limited systems. Journal of Real-Time Image Processing, 10(1):163–174, 2015.

[15] C. Villani. Optimal Transport: Old and New. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2008.

[16] Ali N Yousif, Hassan M Ibrahim, Safaa J Alwan, and Mohammed Sh Majid. Stereo vision development for high performance on stereo systems. International Journal of Nonlinear Analysis and Applications, 13(1):2731–2738, 2022.