On the solution of the photometric stereo problem with unknown lighting

A. Concas, C. Fenu, G. Rodriguez

Department of Mathematics and Computer Science
University of Cagliari, Italy

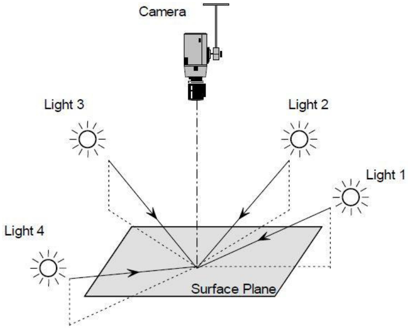
Mathematical Methods for Objects Reconstruction:
from 3D Vision to 3D Printing

INdAM, "Rome", February 10–12, 2021

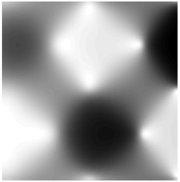
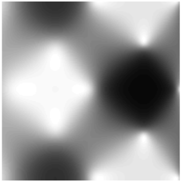
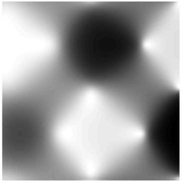
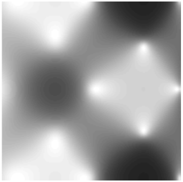
Shape from shading

A typical problem in Computer Vision consists of reconstructing the 3D shape of an object, starting from a set of pictures.

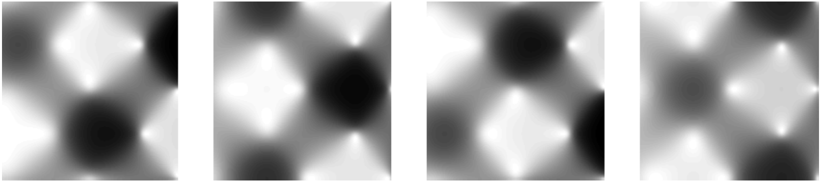
Photometric stereo: the camera and the object are at fixed positions, pictures correspond to different lighting conditions.

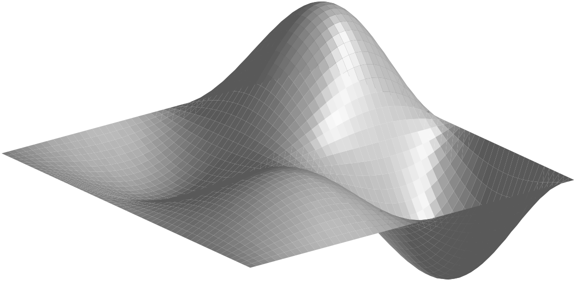


Input

Input & Output



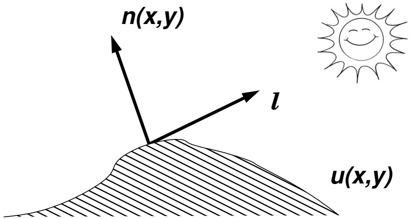


PS in Archaeology

Our research started during a project for documenting petroglyphs found in neolithic tombs called *Domus de Janas* in Sardinia.

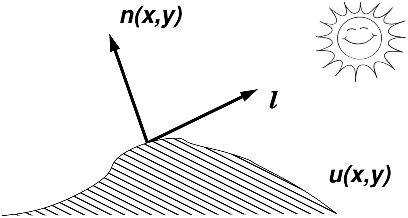
Motivation: obtaining 3D reproductions of *rock art* or artifacts.

The mathematical model: Lambert's cosine law



The *reflectance* at each point is proportional to the cosine of the angle between the normal vector $\mathbf{n}(x, y)$ and the light direction $\boldsymbol{\ell}$.

The mathematical model: Lambert's cosine law



The *reflectance* at each point is proportional to the cosine of the angle between the normal vector $\mathbf{n}(x, y)$ and the light direction $\boldsymbol{\ell}$.

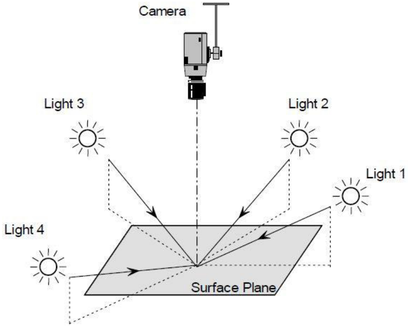
$$\rho(x, y) \cdot \langle \mathbf{n}(x, y), \boldsymbol{\ell} \rangle = \mathcal{I}(x, y)$$

- $\rho(x, y)$ is the *albedo* (depends upon material, paint, etc.)
- for the normal vector, $\|\mathbf{n}(x, y)\| = 1$
- $\|\boldsymbol{\ell}\|$ is proportional to the light intensity
- the surface is represented by $z = u(x, y)$
- $\mathcal{I}(x, y)$ is the observed radiant intensity (prop. to pixel value)

Experimental setting

In the following we will assume that

1. the object is at the origin of a reference system in $\mathbb{R}^3$;
2. the camera is on the z-axis, aiming at the origin;
3. q pictures are available, with light sources at ℓ_t, $t = 1, \ldots, q$;
4. each digital picture $\mathbf{m}_t$ has resolution $r \times s$, with $p = rs$ pixels;
5. the pictures are vectorized (pixels in lexicographic order).



Experimental setting

In the following we will assume that

1. the object is at the origin of a reference system in $\mathbb{R}^3$;
2. the camera is on the z-axis, aiming at the origin;
3. q pictures are available, with light sources at $\boldsymbol{\ell}_t$, $t = 1, \ldots, q$;
4. each digital picture $\mathbf{m}_t$ has resolution $r \times s$, with $p = rs$ pixels;
5. the pictures are vectorized (pixels in lexicographic order).

Physical assumptions:

- the surface is Lambertian;
- there are no self-obstructions, light reflections, or shades;
- the light sources are placed at ∞ distance;
- the camera is sufficiently far from the object to avoid perspective distortions.

None of these assumptions is perfectly met in practice.

There are two main solution approaches

Hamilton–Jacobi formulation ($q \geq 2$)

$$\rho(x,y)\frac{\langle -\nabla u(x,y), \widetilde{\boldsymbol{\ell}}_t\rangle + \ell_{3t}}{\sqrt{1+\|\nabla u(x,y)\|^2}} = \mathcal{I}_t(x,y), \qquad t = 1,\ldots,q,$$

[Kozera, Appl. Math. Comput. 1991], [Mecca, Falcone, SIIMS 2013].

There are two main solution approaches

Hamilton–Jacobi formulation $(q \geq 2)$

$$\rho(x,y)\frac{\langle -\nabla u(x,y), \widetilde{\boldsymbol{\ell}}_t\rangle + \ell_{3t}}{\sqrt{1+\|\nabla u(x,y)\|^2}} = \mathcal{I}_t(x,y), \qquad t = 1,\dots,q,$$

[Kozera, Appl. Math. Comput. 1991], [Mecca, Falcone, SIIMS 2013].

Poisson formulation $(q \geq 3)$

1. Determine (discrete) normal vector field $\mathbf{n}_k$, $k = 1,\dots,p$.
2. Approximate the divergence of the normal field and solve

$$\Delta u(x,y) = f(x,y).$$

For both: Ω is a rectangle, Dirichlet b.c., q lights are known.

Poisson formulation & unknown lighting

In the H–J model the operator to be inverted depends upon the data: the matrix of the discretized problem may be singular or severely ill-conditioned in certain lighting conditions.

On the contrary, in the Poisson approach only the RHS depends on the data.

- Advantage: the computation is decoupled into two simpler problems, allowing for the treatment of **unknown lighting**.
- Drawbacks:
  - it requires a larger number of images (3);
  - to locate lights the SVD of a large dense matrix is needed.

Determining the normal vectors with known lighting - 1

Lambert's law: $\rho(x,y)\cdot\langle\mathbf{n}(x,y),\boldsymbol{\ell}_t\rangle = \mathcal{I}_t(x,y), \quad t=1,\dots,q.$

After discretizing the problem on a regular grid and ordering the pixels lexicographically, we obtain the matrices

$$R = \operatorname{diag}(\rho_1,\dots,\rho_p) \in \mathbb{R}^{p\times p}, \qquad N = [\mathbf{n}_1,\dots,\mathbf{n}_p] \in \mathbb{R}^{3\times p},$$
$$L = [\boldsymbol{\ell}_1,\dots,\boldsymbol{\ell}_q] \in \mathbb{R}^{3\times q}, \qquad M = [\mathbf{m}_1,\dots,\mathbf{m}_q] \in \mathbb{R}^{p\times q},$$

where

$$\rho_k = \rho(x_k,y_k), \quad \mathbf{n}_k = \mathbf{n}(x_k,y_k), \quad k=1,\dots,p,$$

$$\boldsymbol{\ell}_t, \quad \mathbf{m}_t = \operatorname{vec}(\mathcal{I}_t(x_k,y_k)), \quad t=1,\dots,q.$$

Then, Lambert's law can be written as

$$\rho_k\,\mathbf{n}_k^T\boldsymbol{\ell}_t = m_{kt}, \qquad \text{i.e.,} \qquad R\,N^T L = M.$$

Determining the normal vectors with known lighting - 2

Lambert's law: $R\,N^T L = M$

When the light positions are known, we first compute

$$\widetilde{N}^T = ML^\dagger, \qquad \text{where } \widetilde{N}^T := R\,N^T,$$

and then determine the factorization $NR = \widetilde{N}$ by normalizing the columns of $\widetilde{N}$.

Since p (number of pixels) is usually very large, computing R and N requires that $q \geq 3$ and that the ℓ_t vectors are independent.

Finally, we solve the Poisson equation (with Dirichlet b.c.)

$$\Delta u(x,y) = f(x,y),$$

[Dessì, Mannu, R, Tanda, Vanzi, DAACH 2015].

Photometric stereo under unknown lighting

The need for accurate information about the relative position of the lights and the object is a severe limitation of the method.

In practical PS, being able to obtain the lights position opens the possibility of freehand lighting, removing the requirement for accurate positioning of lamps, one of the most difficult issues.

Photometric stereo under unknown lighting

The need for accurate information about the relative position of the lights and the object is a severe limitation of the method.

In practical PS, being able to obtain the lights position opens the possibility of freehand lighting, removing the requirement for accurate positioning of lamps, one of the most difficult issues.

[Hayakawa, J. Opt. Soc. Am. A 1994] proved that the lights position can be determined if at least $q = 6$ images are available, under the usual assumption of a Lambertian surface with lights at ∞ distance.

Determining the lights position - 1

Let the "compact" SVD of the observation matrix be

$$M = U\Sigma V^T,$$

with $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_q)$, $U \in \mathbb{R}^{p\times q}$, and $V \in \mathbb{R}^{q\times q}$.

Since it is expected that $\text{rank}(M) = 3$, we set $W = [\sigma_1\mathbf{u}_1, \sigma_2\mathbf{u}_2, \sigma_3\mathbf{u}_3]^T$ and $Z = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]^T$, so that

$$W^T Z \simeq M.$$

This produces the best rank-3 approximation to the data matrix M with respect to both the Euclidean and the Frobenius norms.

As $3 \leq q \ll p$, when p is very large, the SVD can be approximated by a Lanczos approach [Baglama, Reichel, BIT 2013].

Determining the lights position - 2

Theorem

The normal vectors and the lights position can be uniquely determined from $RN^TL = M$, up to a unitary transformation, only if at least 6 images taken in different lighting conditions are available.

Determining the lights position - 2

Theorem

The normal vectors and the lights position can be uniquely determined from $RN^T L = M$, up to a unitary transformation, only if at least 6 images taken in different lighting conditions are available.

Proof.

We start from the rank-3 factorization $W^T Z = M$, with

$$W = [\mathbf{w}_1, \dots, \mathbf{w}_p] \quad \text{and} \quad Z = [\mathbf{z}_1, \dots, \mathbf{z}_q].$$

Since $\|\ell_t\| = 1$, we seek B such that $\|B\mathbf{z}_t\| = 1$, $t = 1, \dots, q$. This implies solving the system of equations

$$\mathrm{diag}(Z^T G Z) = \mathbf{1},$$

where $G = B^T B$ is symmetric positive definite.

Determining the lights position - 2

$$\operatorname{diag}(Z^T G Z) = \mathbf{1} \quad \Rightarrow \quad \mathbf{z}_t^T G \mathbf{z}_t = \sum_{i,j=1}^{3} z_{it} z_{jt} g_{ij} = 1,$$

for each $t = 1, \ldots, q$, where g_{ij} $(i \leq j)$ are the entries of G.
The system can be rewritten as

$$H\mathbf{g} = \mathbf{1},$$

where $\mathbf{g} = (g_{11}, g_{22}, g_{33}, g_{12}, g_{13}, g_{23})^T$ and $H \in \mathbb{R}^{q \times 6}$ has rows

$$\begin{bmatrix} z_{1t}^2 & z_{2t}^2 & z_{3t}^2 & 2z_{1t}z_{2t} & 2z_{1t}z_{3t} & 2z_{2t}z_{3t} \end{bmatrix}, \qquad t = 1, \ldots, q.$$

The (LS) solution vector $\mathbf{g}$ may be unique only if $q \geq 6$. □

Remark

The matrix B is obtained from G by Cholesky factorization.
Then, $\widetilde{N} = B^{-T} W$, $L = BZ$.

The solution is not unique!

Only the relative position of lights is determined.
[Belhumeur, Kriegman, Yuille, Int. J. Comput. Vis. 1999]

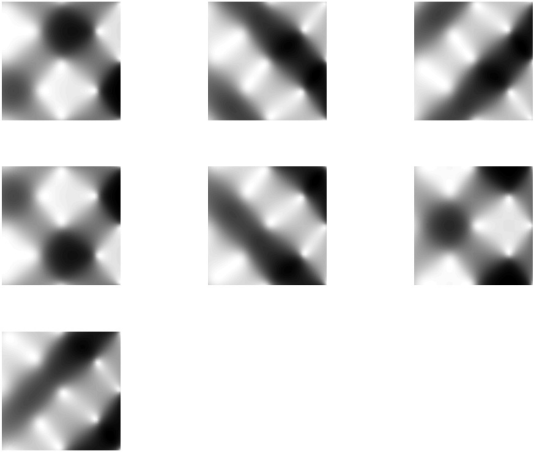
Given an admissible rank-3 factorization

$$\widetilde{N}^T L = M,$$

any matrix pair $(Q\widetilde{N}, QL)$, with $Q \in \mathbb{R}^{3\times 3}$ unitary, is a solution as well.

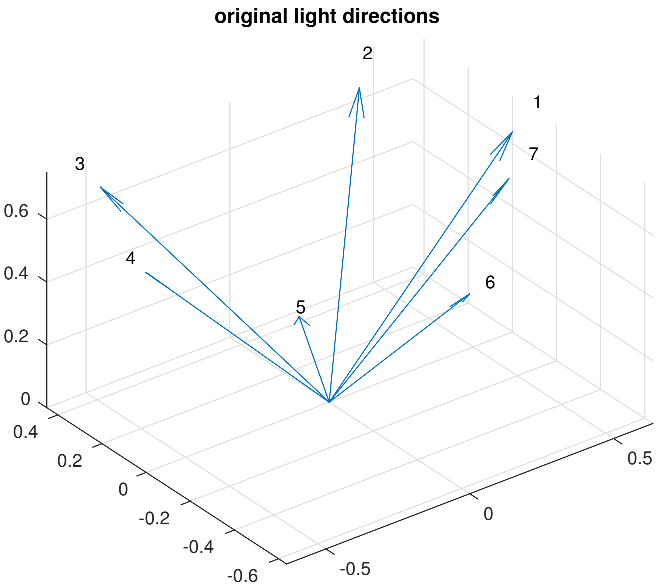
- The system object-camera-lights may result rotated in the reconstruction, perhaps with axes inversions.
- A suitable rotation is essential, if the object is represented by an explicit function $z = u(x, y)$. It can be determined if a suitable shooting technique is adopted.

A numerical experiment

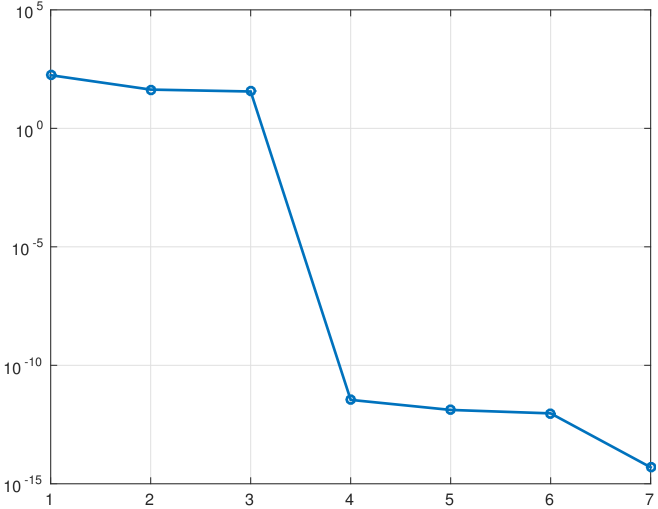


synthetic data set (100 $\times$ 100 pixels)

A numerical experiment

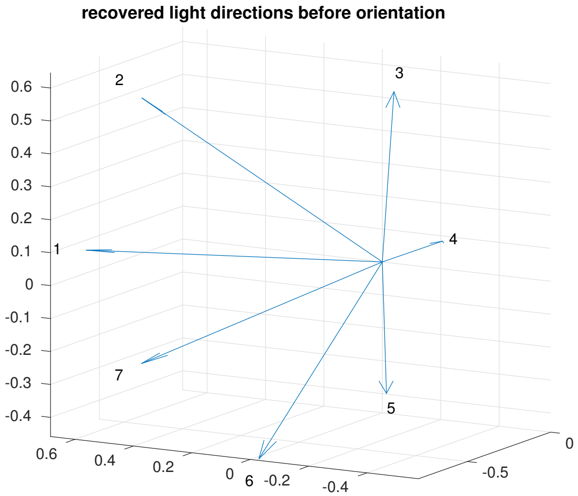


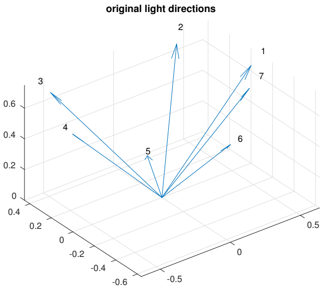
A numerical experiment



singular values of the data matrix M

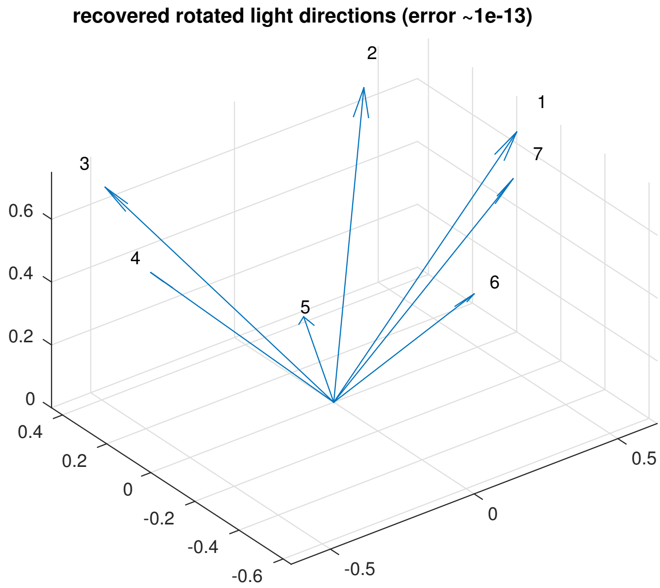
A numerical experiment



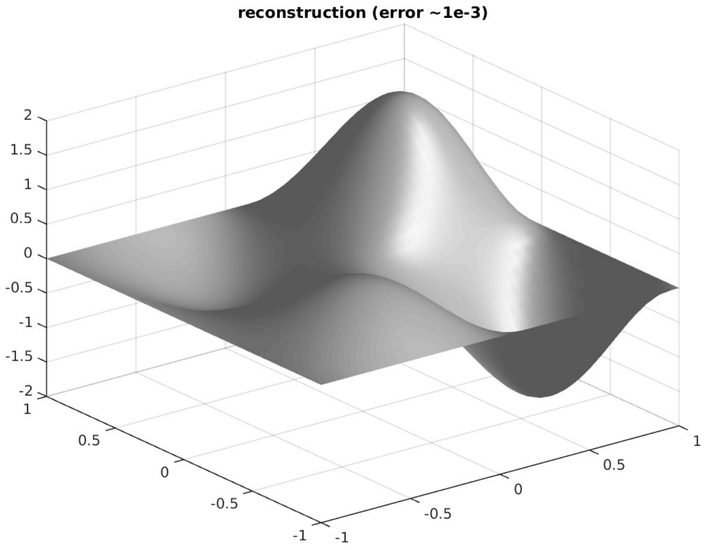


the Hayakawa procedure produces an axis inversion

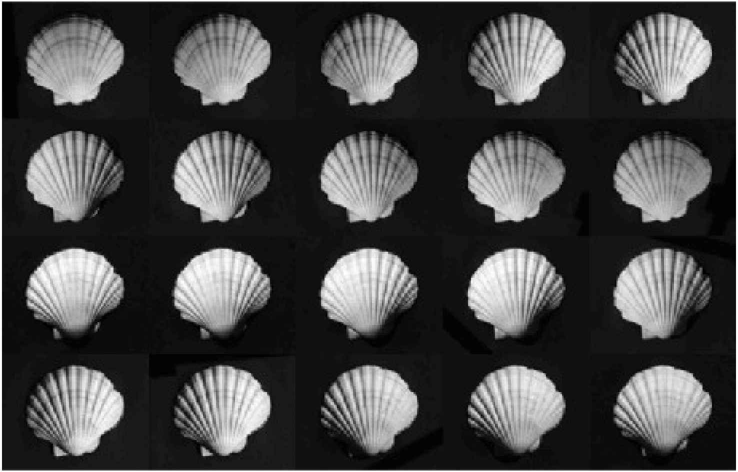
A numerical experiment



A numerical experiment

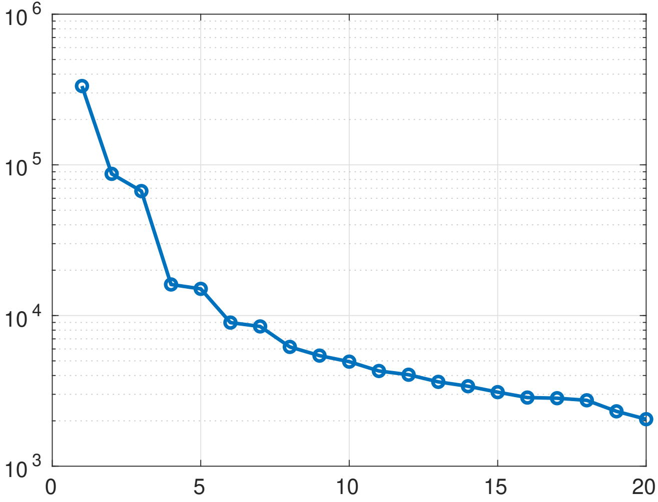


An experimental data set with “quasi-ideal” lighting



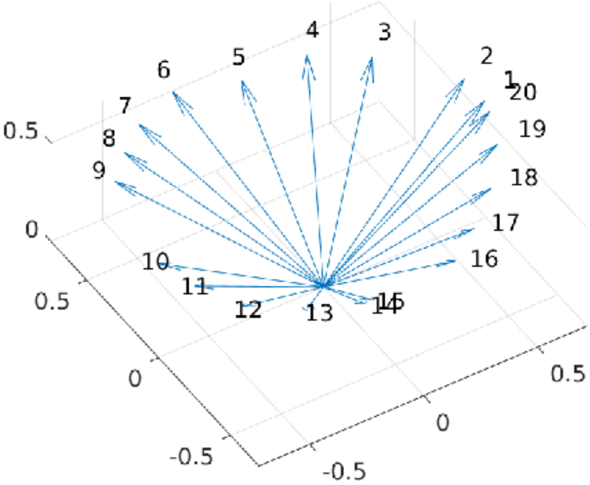
SHELL data set: 20 pictures (885 $\times$ 705), sun light

An experimental data set with "quasi-ideal" lighting

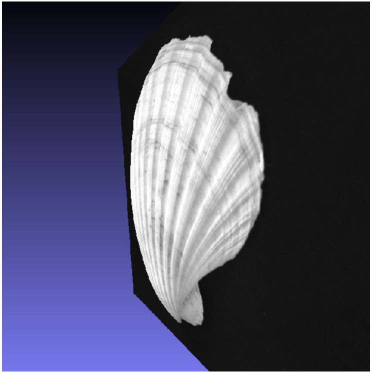


singular values of the data matrix M

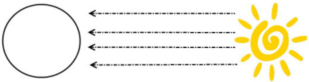
An experimental data set with “quasi-ideal” lighting



An experimental data set with “quasi-ideal” lighting



Working in unideal light conditions

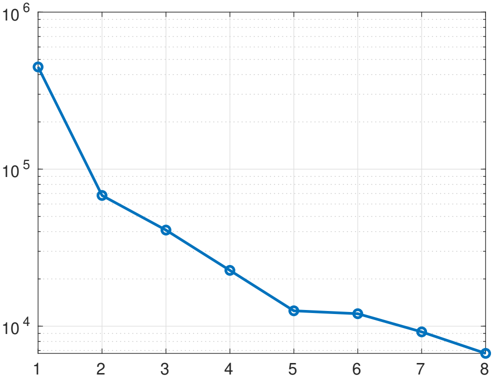


Real light sources are often far from being ideal

1. they may be placed close to the surface (especially in narrow locations, like caves or excavation sites)
2. they are differently attenuated at different points of the object, according to distance

We must also keep into account the deformation caused by the finite (often small) distance between the camera and the object, though it might be corrected by camera calibration techniques.

How does the algorithm perform in this light conditions?



Ptolemaic stela (39 B.C.) from the Museo Egizio (Torino), 8 pictures

The STELA data set

Ptolemaic stela (39 B.C.) from the Museo Egizio (Torino), 8 pictures

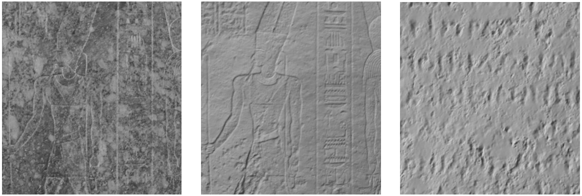
Synthetic data with lights at fixed distance

distance	σ_3/σ_4	E_{lights}	E_{surface}
∞	8.74e+14	1.40e-15	1.08e-03
1000	6.21e+03	4.62e-05	2.53e-03
100	6.21e+02	4.67e-04	2.33e-02
10	6.18e+01	5.20e-03	2.57e-01
1	5.21e+00	9.66e-02	1.25

computer generated images
unit for distance is object width
σ_3/σ_4 represents "closeness to rank 3"
errors are relative in the Frobenius norm

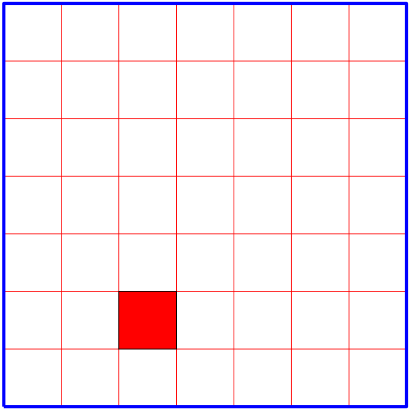
How to deal with close lights?

An alternative to adopting a more complex mathematical model, we are working on an approximation approach.



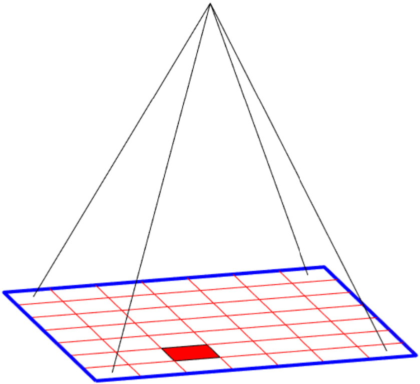
Numerical tests on small portions of a large object suggest that when the lights cannot be moved away from the object, accuracy can be improved by reducing the size of the domain.

Solution: domain decomposition



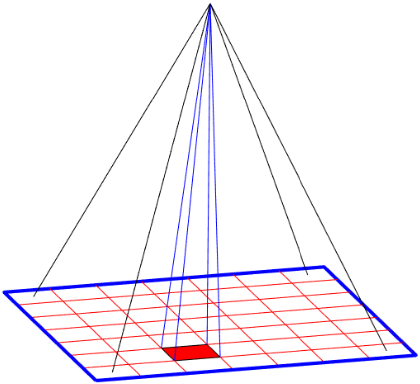
the process can be applied to a subdomain

Solution: domain decomposition



light rays are divergent for the whole domain,

Solution: domain decomposition



light rays are divergent for the whole domain,
but they are almost parallel for a small sudomain

Domain decomposition and parallelization

Decomposing the problem in subdomains may be useful at various levels:

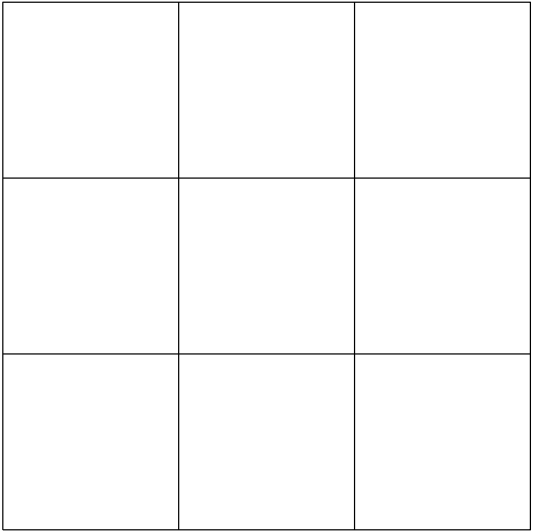
- Parallelize the algorithm and optimize performance (even with lights at ∞).
- Reduce the distortion deriving from non-ideal conditions, i.e., with lights at a finite distance.
- Schwharz domain decomposition for PDE's.
- Localize lights at finite distance and resort to a different, more complex, model.

Domain decomposition and parallelization

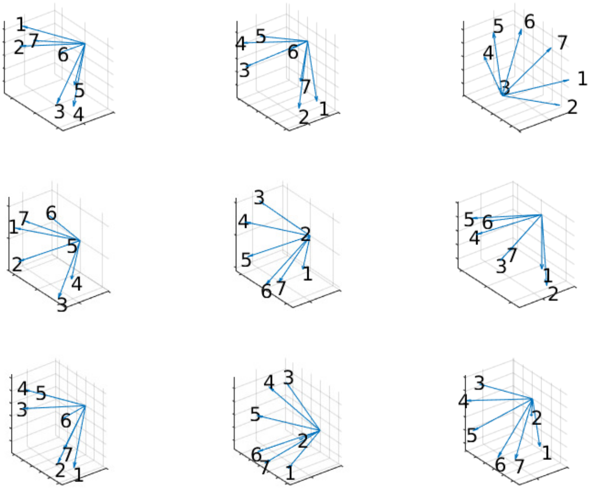
Decomposing the problem in subdomains may be useful at various levels:

- Parallelize the algorithm and optimize performance (even with lights at ∞).
- Reduce the distortion deriving from non-ideal conditions, i.e., with lights at a finite distance.
- Schwharz domain decomposition for PDE's.
- Localize lights at finite distance and resort to a different, more complex, model.

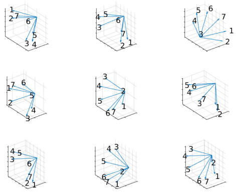
The computation cannot be fully parallelized



The computation cannot be fully parallelized

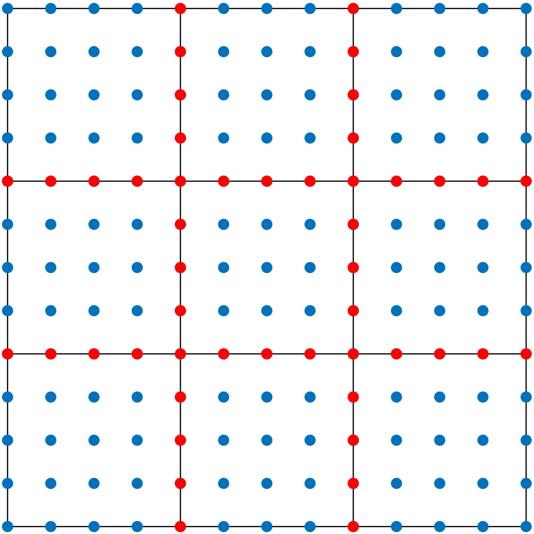


The computation cannot be fully parallelized

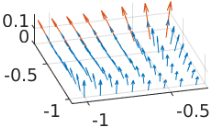


- If lights are at ∞, their direction can be used as a reference for determining a coherent rotation.
- This cannot be done if lights are at a finite distance, as they are seen from different points of view.

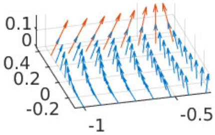
Computing the normals on the subdomains

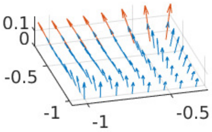


Computing the normals on the subdomains

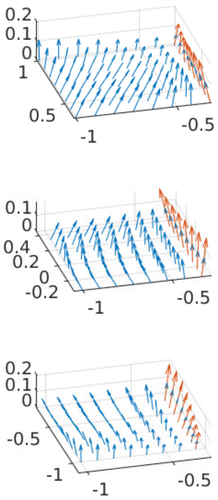


Computing the normals on the subdomains

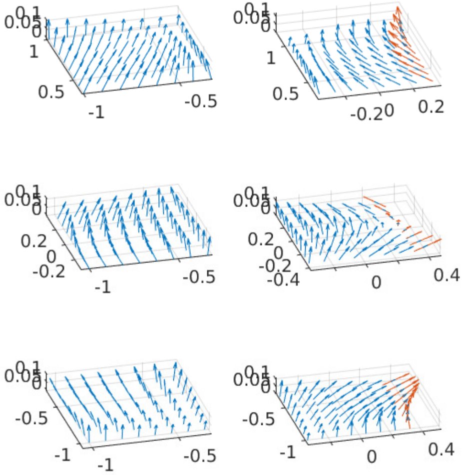




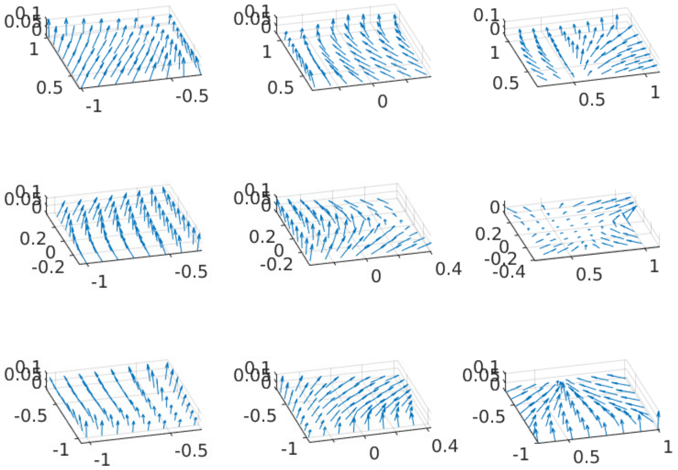
Computing the normals on the subdomains



Computing the normals on the subdomains



Computing the normals on the subdomains



Determining the correct rotation

Let $N_1 = N_{1,1}^{\text{north}}$ and $N_2 = N_{1,2}^{\text{south}}$ denote the "north" and "south" boundary normal vectors for the first two domains.

We seek a unitary matrix Q which solves the optimization problem

$$\min_{Q^T Q = I} \|N_1 - QN_2\|_F.$$

Determining the correct rotation

Let $N_1 = N_{1,1}^{\text{north}}$ and $N_2 = N_{1,2}^{\text{south}}$ denote the "north" and "south" boundary normal vectors for the first two domains.

We seek a unitary matrix Q which solves the optimization problem

$$\min_{Q^T Q = I} \|N_1 - QN_2\|_F.$$

The solution is obtained by computing the SVD

$$N_2 N_1^T = U \Sigma V^T,$$

and setting $Q = VU^T$ [Higham, book chapter 1989].

Computing Q is really not a problem as $N_2 N_1^T \in \mathbb{R}^{3\times3}$.

Software and testing *"under construction"*

- A Matlab suite for PS3D with unknown lighting has been developed and it will be available soon.

- We have a prototypal version of the domain decomposition software, but it will take time to optimize it.

- On small size synthetic datasets (10^4–10^5 pixels) we observed a 15% reduction in the reconstruction error with 3×3 and a 30% reduction in the reconstruction error with 5×5 subdomain grids.

Finally, . . .

thank you for your attention!