Chapter 12
Photometric 3D-reconstruction

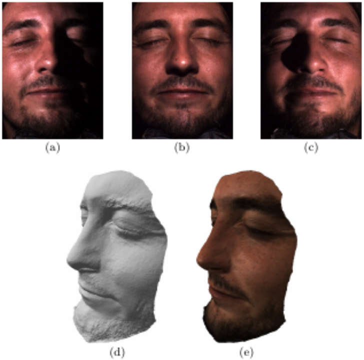
Computer Vision I: Variational Methods

Winter 2017/18

Dr. Yvain QUÉAU
Chair for Computer Vision and Pattern Recognition
Departments of Informatics & Mathematics
Technical University of Munich

3D-scanning

3D-scanning = estimation of shape *and* color



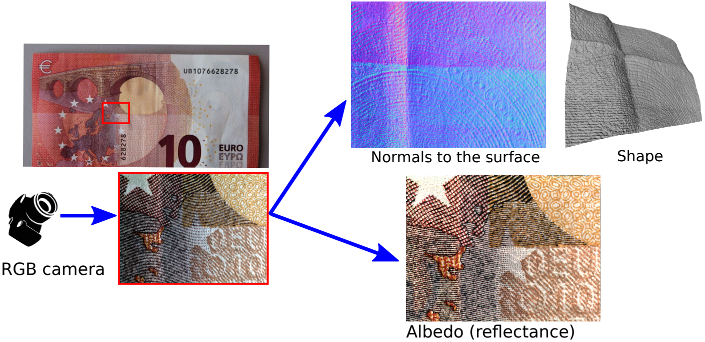
[Quéau et al., 2017b, (JMIV)]

- Geometric techniques (e.g. SfM, MVS) based on multi-view consistency only recover shape (no color estimation)
- Photometric techiques (e.g., shape-from-shading and photometric stereo), which are based on inverting the image formation model, recover both.

Photometric techniques

Photometric techniques constitute the first choice for the recovery of very thin structures.

Example: close-up on a 10 euros banknote



[Quéau, 2015, (PhD thesis)]

Photometric 3D-reconstruction =

Shape analysis through *luminous quantities* (photo) *measurement* (metric), by reverting the image formation process.

Photometric techniques

Photometric techniques constitute the first choice for the recovery of very thin structures.

Example: close-up on a 5 euros banknote

[Quéau et al., 2017g, (QCAV)]

Photometric techniques

Photometric techniques constitute the first choice for the recovery of very thin structures.

Example: close-up on a 50 cents euros coin

[Quéau et al., 2016a, (CVPR)]

Photometric techniques

Photometric techniques constitute the first choice for the recovery of very thin structures.

Example: close-up on human skin

Overview

1. **Basics of photometric 3D-reconstruction**

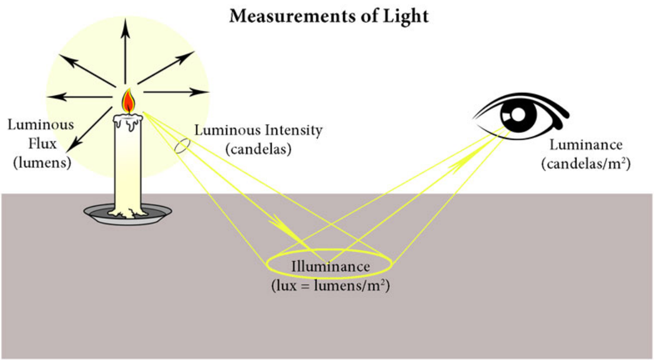
2. **Shape-from-shading**

3. **Photometric stereo**

Overview

1. **Basics of photometric 3D-reconstruction**
2. **Shape-from-shading**
3. **Photometric stereo**

A few photometric definitions

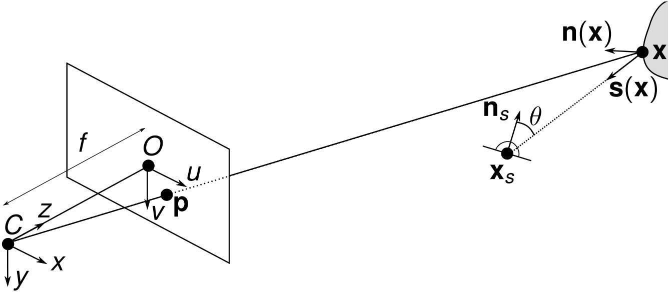


(source: `http://rsagencies.co.za`)

- luminance (radiance) $=$ perceived luminosity - cd.m^{-2} $=$ lm.m^{-2}.sr^{-1} (W.m^{-2}.sr^{-1})
- illuminance (irradiance): luminous flux received by a surface per unit area - lx $=$ lm.m^{-2} (W.m^{-2})
- reflectance: proportion of incident light which is reflected by a surface (unit-less)

Digital cameras as luminance measurement devices

Consider a pixel p conjugate to a 3D-point x on the surface.

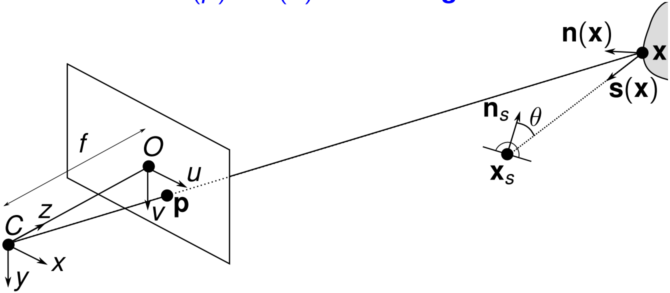


- Digital cameras essentially count the number of photons crossing the cell i.e., the pixel p
- $I(p) \approx \frac{t}{g} E_s(p)$ with t the exposure time, g the gain and $E_s(p)$ the sensor illuminance at pixel p
- $E_s(p) \approx L(x)$ with $L(x)$ the luminance of the point x on the surface

$\rightarrow$ The brightness in p is proportional to the luminance in x:

$$I(p) \approx L(x)$$

How is formula $I(p) \approx L(x)$ interesting?



Further consider that incident light in x is coming from all directions ω on the visible hemisphere $\mathcal{H}(x)$. Then:

$$I(p) \approx L(x) = \iint_{\omega \in \mathcal{H}(x)} \rho(x) \phi_\omega \max\{0, s_\omega \cdot n(x)\} \mathrm{d}\omega$$

with ϕ_ω the light intensity (in cd) in direction ω, s_ω the unit-length lighting direction, $\rho(x)$ the reflectance in point x **(i.e., the surface material)** and $n(x)$ the normal to the surface **(i.e., the surface material)**

$\rightarrow$ By inverting this image formation model, one may reconstruct the surface shape and its material (color, roughness, etc.)

Inverting the image formation model

In general, we need to simplify

$$I(p) \approx L(x) = \int_{\omega \in \mathcal{H}(x)} \rho(x, \ldots) \phi_\omega \max\{0, s_\omega \cdot n(x)\} \mathrm{d}\omega$$

because of:

- the integral over all lighting directions (partial solution: spherical harmonics, cf. talks of Bjoern and Robert)
- the complexity of the reflectance function ρ, which generally depends on the angle between incident light and viewing direction (BRDF, for bidirectional reflectance function, see for instance lectures on computer graphics)

(remark: considering multi-channel images adds even more complexity - be careful if you read that such extensions are straightforward ;-))

A simple, yet effective image formation model

- Assuming a **single light source**, the integrand is a Dirac:

$$\int_{\omega\in\mathcal{H}(x)} \rho(x)\phi_\omega \max\{0, s_\omega\cdot n(x)\}\mathrm{d}\omega = \rho(x,\dots)\phi(x)\max\{0, s(x)\cdot n(x)\}$$

with $s(x)$ the unit-length direction of the the luminous flux reaching x and $\phi(x)$ its density (in cd.m^{-2}).

- Assume a **Lambertian surface: luminance is independent from the viewing angle**. Corollary:

$$\rho(x,\dots) := \frac{\rho(x)}{\pi}$$

with $\rho(x) \geq 0$ a scalar function called the **albedo** ($0 =$ black, $1 =$ white).

And thus:

$$I(p) \approx \rho(x)\phi(x)\max\{0, s(x)\cdot n(x)\}$$

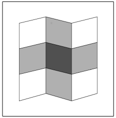
(often **abusively** referred to as Lambert's law)

Inverting the simplified image formation model

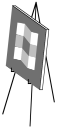
Considering the relationship:

$$I(p) \approx \rho(x)\phi(x) \max\{0, s(x) \cdot n(x)\}$$

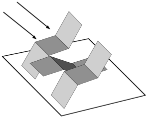
one may want to estimate shape (n), reflectance (ρ) and lighting (ϕ and s). **In general, this is impossible.**



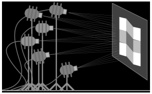
An image



Painter's explanation



Sculptor's explanation



Lighting designer's explanation

Impossible to tell **reflectance** from **shape** and **lighting** (images from [Adelson and Pentland, 1996, (Perception as Bayesian inference)])

Overview

1. Basics of photometric 3D-reconstruction

2. **Shape-from-shading**

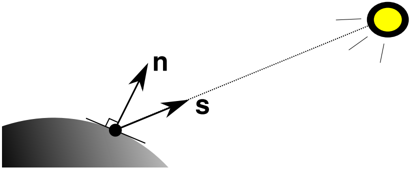
3. Photometric stereo

Shape-from-shading

Shape-from-shading [Horn, 1970] is the task of estimating shape, assuming known lighting and reflectance. Classically one further simplifies the previous image formation model

$$\underbrace{I(p)}_{\text{Luminance}} = \underbrace{\rho(x)}_{\text{Albedo}} \underbrace{\phi(x)}_{\text{Intensity}} \underbrace{\max\{0, \mathbf{s}(x) \cdot \mathbf{n}(x)\}}_{\text{shading}}$$

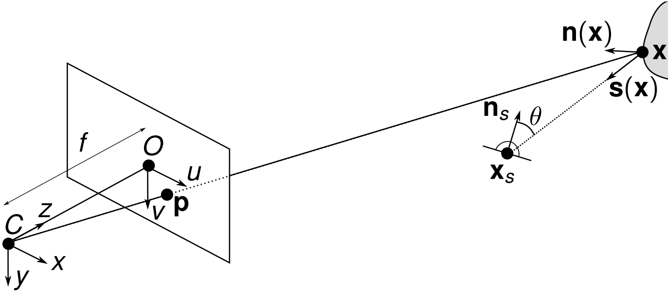
where the *shading* (i.e., the surface illuminance) is the the dot product between the unit-length normal and lighting directions:



- $\mathbf{s} \in \mathbb{S}^2 \subset \mathbb{R}^3$: lighting direction
- $\mathbf{n} \in \mathbb{S}^2 \subset \mathbb{R}^3$: normale to the surface (shape)

Classical assumptions of shape-from-shading

Consider a 3D-frame attached to the camera:



and further assume:

- White surface: $\rho(x) := 1$
- Distant point light source: $\phi(x) := \phi > 0$,
  $\mathbf{s}(x) := \mathbf{s} \in \mathbb{S}^2 \subset \mathbb{R}^3$
- Calibrated source: $\phi := 1$
- Frontal lighting: $\mathbf{s} := [0, 0, -1]^\top$
- Orthographic camera: $\mathbf{n}(x) := \frac{1}{\sqrt{|\nabla z(p)|^2+1}} \begin{bmatrix} \nabla z(p) \\ -1 \end{bmatrix}$,

  (z: **depth map**)

Eikonal SfS model

Plugging $\rho(x) := 1$, $\mathbf{s}(x) := [0, 0, -1]^\top$, $\phi(x) := 1$, and $\mathbf{n}(x) := \frac{1}{\sqrt{|\nabla z(p)|^2+1}} \begin{bmatrix} \nabla z(p) \\ -1 \end{bmatrix}$ in our model

$$\underbrace{I(p)}_{\text{Luminance}} = \underbrace{\rho(x)}_{\text{Albedo}} \underbrace{\phi(x)}_{\text{Intensity}} \underbrace{\max\{0, \mathbf{s}(x) \cdot \mathbf{n}(x)\}}_{\text{shading}},$$

we obtain a nonlinear PDE in z over the image domain Ω:

$$I(p) = \frac{1}{\sqrt{|\nabla z(p)|^2 + 1}}, \; p \in \Omega \subset \mathbb{R}^2.$$

This is directly related to the celebrated **eikonal equation** $|\nabla z| = f$. The unique maximum viscosity solution of SfS can thus be determined by solving the eikonal equation [Lions et al., 1993].

On uniqueness in shape-from-shading

Despite the extremely restrictive assumptions listed above, SfS remains ambiguous. For instance, the two following results are equally valid explanations of the Lena image under such assumptions:

Maximum viscosity solution [Cristiani and Falcone, 2007]

Variational solution [Quéau et al., 2017f, (EMMCVPR)]

On uniqueness in shape-from-shading

Some solutions to disambiguate shape-from-shading

0 consider an RGB image with three colored light sources [Kontsevich et al., 1994]

0 relax the distant light assumption [Prados and Faugeras, 2005]

1. Add priors within a **variational framework** [Horn and Brooks, 1986]
2. Consider multiple images under varying lighting: **photometric stereo** [Woodham, 1980]

The variational approach to shape-from-shading

[Horn and Brooks, 1986] decouple local and global shape estimations, introducing an integrability prior

1) Estimate an **integrable** (curl-free) gradient field $(p, q) := \nabla z$:

$$\min_{p,q} \iint \left(I - \frac{1}{\sqrt{p^2 + q^2 + 1}} \right)^2 + \lambda (\partial_y p - \partial_x q)^2 \, \mathrm{d}x \mathrm{d}y$$

(nonlinear global optimization problem: hard to solve)

2) **Integrate** this field into a depth map:

$$\min_{z} \iint \|(p, q) - \nabla z\|^2 \, \mathrm{d}x \mathrm{d}y$$

(not as straightforward as it seems to be [Quéau et al., 2017c, Quéau et al., 2017d, (JMIV)])

Some recent developments [Quéau et al., 2017f, (EMMCVPR)]

(p, q) is conservative *by construction* $\rightarrow$ joint local and global (integrated) estimation:

$$\min_{p,q,z} \iint \left(I - \frac{1}{\sqrt{p^2 + q^2 + 1}} \right)^2 \mathrm{d}x\mathrm{d}y$$
$$\text{s.t. } (p, q) = \nabla z$$

One can add smoothness and shape priors, if needed:

$$\min_{p,q,z} \iint \lambda \left(I - \frac{1}{\sqrt{p^2 + q^2 + 1}} \right)^2 + \nu \sqrt{1 + p^2 + q^2}\, \mathrm{d}x\mathrm{d}y$$
$$+ \iint \mu \left(Kz - z^0 \right)^2 \mathrm{d}x\mathrm{d}y$$
$$\text{s.t. } (p, q) = \nabla z$$

Playing with parameters and integration domains, one may achieve shape-from-shading, depth denoising and inpainting, depth refinement and super-resolution, etc. Numerical solution using ADMM.

Example of result (without regularisation)

Other straightforward extensions include RGB data, spherical harmonics lighting, perspective camera, etc. Example of SfS results from [Quéau et al., 2017f] applied to the dataset from [Han et al., 2013]:

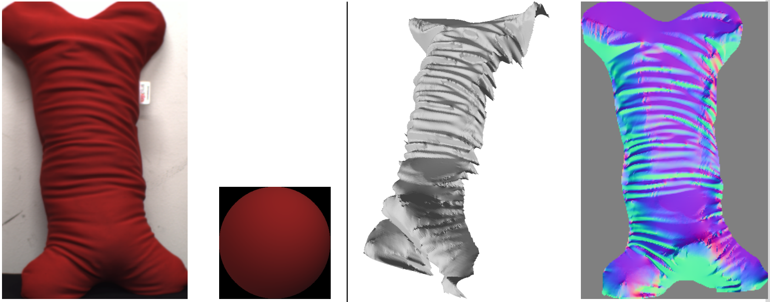
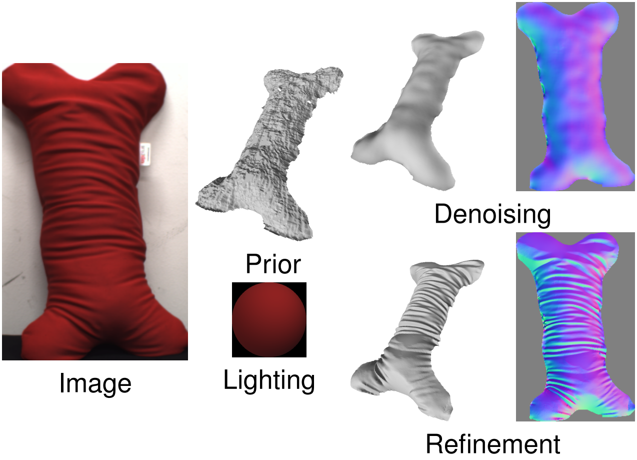


Image | Lighting | 3D-reconstruction

Example of result (with regularisation)



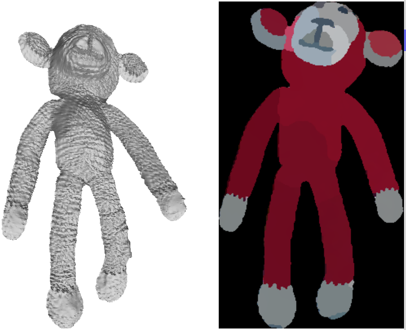
Matlab code available:

`https://github.com/yqueau/shape_from_shading`

Depth super-resolution for RGB-D sensors

By including lighting (spherical harmonics) and albedo (Mumford-Shah prior), the previous model can achieve state-of-the-art single-shot super-resolution of depth maps for RGBD sensors [Häfner et al., 2018].

Data: LR depth
+ HR image



HR depth
+ HR reflectance

(cf. seminar of Bjoern Haefner)

Overview

1 Basics of photometric 3D-reconstruction

2 Shape-from-shading

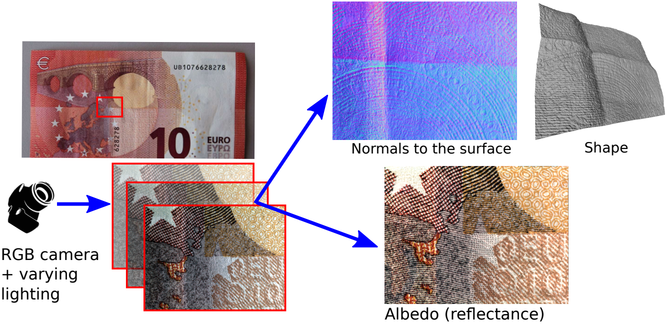
3 **Photometric stereo**

Photometric stereo

When no prior knowledge is available, it is hopeless to achieve a reasonable 3D-reconstruction based on SfS, due to its ill-posedness.

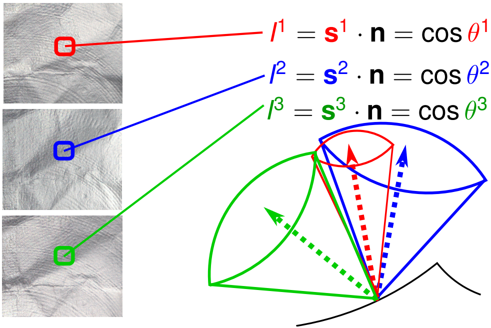
Photometric stereo [Woodham, 1980] is an extension of SfS which considers multiple images of the surface, taken from the **same viewing angle** but under **varying lighting**.

Besides its ability to unambiguously recover thin geometric structures, photometric stereo is the only computer vision technique which is able to **estimate reflectance**.



[Quéau, 2015, PhD thesis]

Why photometric stereo disambiguates SfS



- $m = 1$ image (SfS), known albedo: infintely many possible normals in each point
- $m = 2$ images, known albedo: two possible normals [Quéau et al., 2017e, (IVC)]
- $m \geq 3$: unique approximate solution - problem is over-constrained and thus albedo can be estimated

Photometric stereo as a linear system of equations

Recall our initial model:

$$\underbrace{I(p)}_{\text{Luminance}} = \underbrace{\rho(x)}_{\text{Albedo}}\ \underbrace{\phi(x)}_{\text{Intensity}}\ \underbrace{\max\{0, \mathbf{s}(x)\cdot\mathbf{n}(x)\}}_{\text{shading}}$$

Consider $m \geq 3$ images I^i under known, varying lighting $(\phi^i(x), \mathbf{s}^i(x))$. Then one needs to solve, in each pixel p, the system of equations in $(\rho(x) > 0, \mathbf{n}(x) \in \mathbb{S}^2)$:

$$I^i(p) = \rho(x)\phi^i(x)\max\left\{0, \mathbf{s}^i(x)\cdot\mathbf{n}(x)\right\}, \quad i \in \{1, \dots, m\}$$

Assuming distant light sources ($s^i(x) = \phi^i(x)\mathbf{s}^i(x) := s^i \in \mathbb{R}^3$), neglecting self-shadows ($\max\{a, 0\} := a$), and plugging together the unknowns into

$$m(x) := \rho(x)\mathbf{n}(x),$$

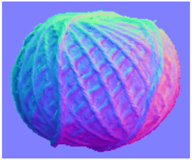
then we obtain a linear system of equations in $m(x)$:

$$I^i(p) = s^i \cdot m(x), \quad i \in \{1, \dots, m\}$$

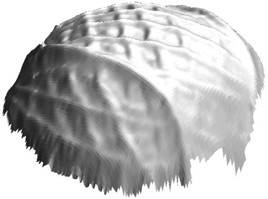
which we can solve and then deduce $\rho(x) := |m(x)| > 0$ and $\mathbf{n}(x) := \frac{m(x)}{|m(x)|}$.

Two-stages resolution of photometric stereo

→ (1)



↓ (2)



(1) Estimation of surface normals (and albedo)

(2) Integration of normals [Quéau et al., 2017c, (JMIV)]

This approach can even be extended to unknown lighting, for instance by using TV regularisation [Quéau et al., 2015].

Basic photometric stereo strategy

Data

- lighting vectors
  $s^1, s^2, s^3, \cdots \in \mathbb{R}^3$
- brightness
  $I^1, I^2, I^3, \cdots > 0$

1. $\min\limits_{m \in \mathbb{R}^3} \sum_{i=1}^{m} \Phi\left(\left|s^i \cdot m - I^i\right|\right)$
   with Φ some robust estimator
2. $\rho = \|m\|$
3. $\mathbf{n} = \dfrac{m}{\|m\|}$
4. $\mathbf{n} \to z$ (integration)


I^1


I^2

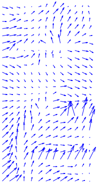

I^3


Albedo ρ

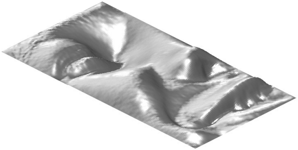

Normals $\mathbf{n}$


Depth z

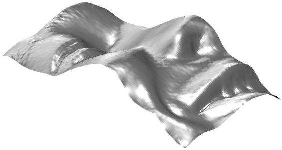
Integrating the normals (1)



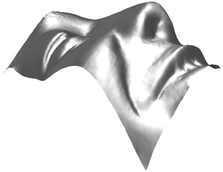
Normals



Integration (Dirichlet)



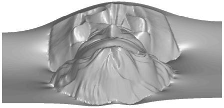
Integration (periodic)



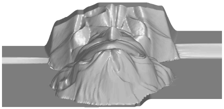
Integration (natural Neumann)

Integrating the normals (2)

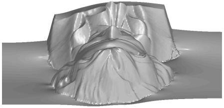
$$\min_{z} \iint_{x\in\Omega} \Phi\left(\left\| \nabla z - [-m_1/m_3,\, -m_2/m_3]^\top \right\| \right) \mathrm{d}x$$



Φ = least-squares



Φ = norm 1



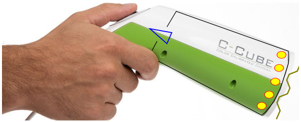
Φ non-convex

- Least squares: easy / over-smoothing
- Norm 1: convex / staircasing
- Non-convex: harder / better (not faster, but stronger ;))

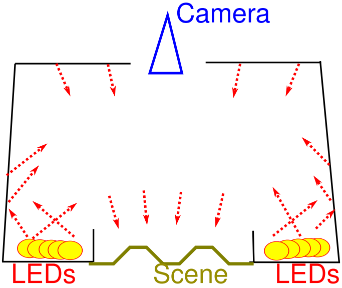
[Quéau et al., 2017d, (JMIV)]

Experimental setup

(joint work with the Pixience company and Toulouse Tech Transfer)



Dermoscope with LEDs



Schematic representation

[Quéau et al., 2017g, QCAV]

3D-reconstruction of metallic coins

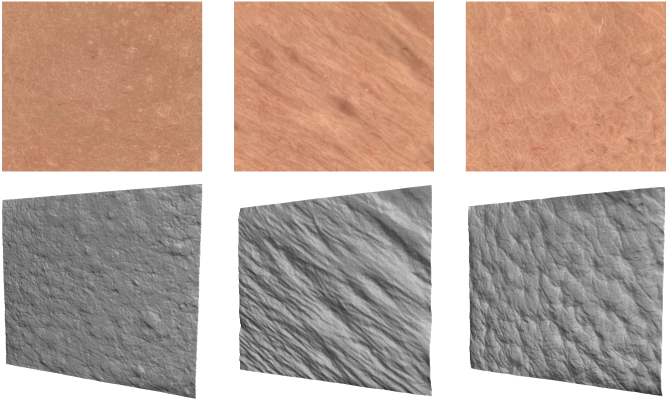
1 euro (Italie)

50 cents (Espagne)

1 yuan (Chine)

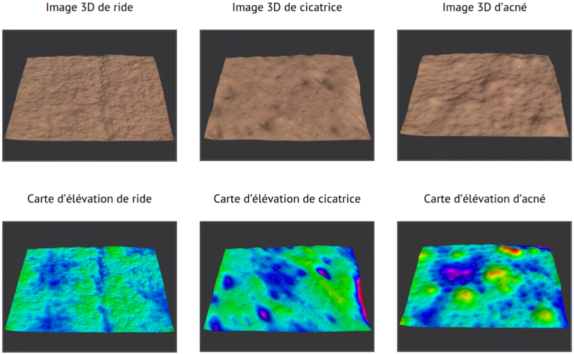
3D-reconstructions [Quéau et al., 2016a, (CVPR)]

3D-reconstructions of synthetic skin samples



3D-reconstructions [Quéau et al., 2016a, (CVPR)]

3D-reconstructions of the human skin



Source : `http://www.pixience.com/produits-2/c-cube-recherche-clinique/module-3d/`

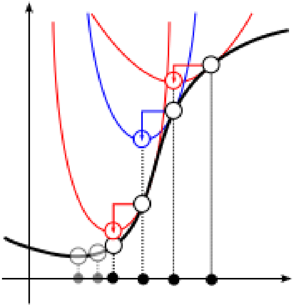
Integrated variational approach

Direct estimation of depth $z : \Omega \to \mathbb{R}$ and albedo $\rho : \Omega \to \mathbb{R}$ solutions of $I^i = \rho s^i \cdot \frac{[\nabla z^\top, -1]^\top}{\sqrt{|\nabla z|^2+1}}$ [Quéau et al., 2017i, (CVPR)]

Linearisation $\rho := \frac{\rho}{\sqrt{|\nabla z|^2+1}}$, then

$$\min_{z,\rho} \sum_{i=1}^{m} \iint_{x \in \Omega} \Phi \left(\left| \rho(x) \mathbf{s}^i \cdot \left[\nabla z(x)^\top, -1 \right]^\top - I^i(x) \right| \right) \mathrm{d}x$$

Non-convex problem: optimisation by alternating reweighted least-squares (alternating MM) [Quéau et al., 2017b, Quéau et al., 2017i, Quéau et al., 2017h, Mélou et al., 2017]

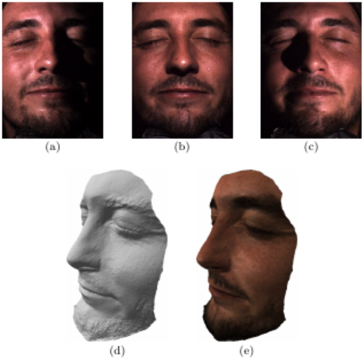


Straightforward extensions of the integrated variational approach

$$\min_{z,\rho} \sum_{i=1}^{m} \iint_{x \in \Omega} \Phi\left(\left| \rho(x) s^i \cdot \left[\nabla z(x)^\top, -1 \right]^\top - I^i(x) \right| \right) \mathrm{d}x$$

- Perspective projection [Quéau et al., 2017i, (CVPR)]
- Uncalibrated PS [Quéau et al., 2017i, (CVPR)]: just include lighting vectors s^i in the unknowns – there is still a unique solution under perspective projection
- Color images and reflectance [Quéau et al., 2017b, (JMIV)]
- Nearby, anisotropic light sources [Quéau et al., 2017b, Quéau et al., 2017h, (SSVM + JMIV)] (s^i becomes a nonlinear function of z)
- Depth prior from RGBD sensor [Peng et al., 2017, Quéau et al., 2017a, ICCV + TS]

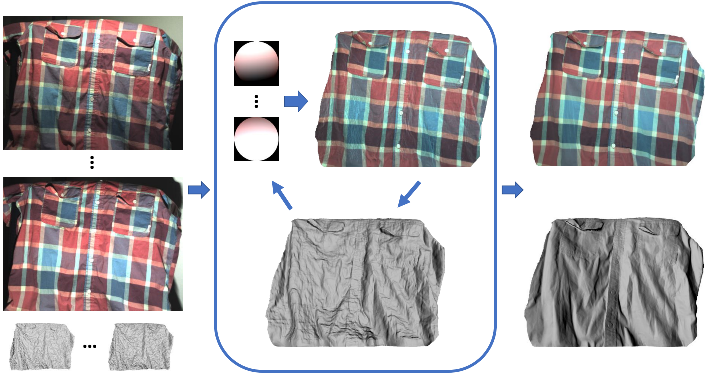
Example: LED-based photometric stereo [Quéau et al., 2017b, (JMIV)]



Codes: `https://github.com/yqueau/near_ps`

Example: uncalibrated photometric stereo-based depth super-resolution [Peng et al., 2017, (ICCV)]

(cf. seminar of Bjoern Haefner)



Codes: `https://github.com/pengsongyou/SRmeetsPS`

Just for finishing this series of lectures on a cool (who said useless ?) application

We can use PS to scan people while they are watching their holidays pictures [Quéau et al., 2016b]

Bibliography I

Adelson, E. H. and Pentland, A. P. (1996).
The perception of shading and reflectance.
In *Perception as Bayesian inference*, pages 409–423. Cambridge University Press.

Cristiani, E. and Falcone, M. (2007).
Fast Semi-Lagrangian Schemes for the Eikonal Equation and Applications.
SIAM Journal on Numerical Analysis, 45(5):1979–2011.

Häfner, B., Quéau, Y., Möllenhof, T., and Cremers, D. (2018).
Fight ill-posedness with ill-posedness: single-shot variational depth super-resolution from shading.
(submitted to CVPR).

Han, Y., Lee, J.-Y., and Kweon, I.-S. (2013).
High quality shape from a single RGB-D image under uncalibrated natural illumination.
In *International Conference on Computer Vision*.

Horn, B. K. P. (1970).
Shape from Shading: a Method for Obtaining the Shape of a Smooth Opaque Object from One View.
Ph.d thesis, MIT, Cambridge, USA.

Horn, B. K. P. and Brooks, M. J. (1986).
The variational approach to shape from shading.
Computer Vision, Graphics, and Image Processing, 33(2):174–208.

Kontsevich, L. L., Petrov, A. P., and Vergelskaya, I. S. (1994).
Reconstruction of shape from shading in color images.
Journal of the Optical Society of America A, 11(3):1047–1052.

Lions, P. L., Rouy, E., and Tourin, A. (1993).
Shape-from-shading, viscosity solutions and edges.
Numerische Mathematik, 64(1):323–353.

Mélou, J., Quéau, Y., Durou, J.-D., Castan, F., and Cremers, D. (2017).
Variational Reflectance Estimation from Multi-view Images.
Journal of Mathematical Imaging and Vision.
(in revision).

Bibliography II

Peng, S., Häfner, B., Quéau, Y., and Cremers, D. (2017).
Depth Super-Resolution Meets Uncalibrated Photometric Stereo.
In *IEEE International Conference on Computer Vision Workshops (ICCVW)*, pages 2961–2968.

Prados, E. and Faugeras, O. (2005).
Shape from shading: a well-posed problem?
In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 870–877.

Quéau, Y. (2015).
Reconstruction tridimensionnelle par stéréophotométrie.
Thèse de doctorat, Université de Toulouse.

Quéau, Y., Durix, B., Lucas, T., Boumaza, J., Durou, J.-D., and Lauze, F. (2017a).
Fusion de données RVB-D par stéréophotométrie colorée.
Traitement du Signal.
(in French, to appear).

Quéau, Y., Durix, B., Wu, T., Cremers, D., Lauze, F., and Durou, J.-D. (2017b).
LED-based Photometric Stereo: Modeling, Calibration and Numerical Solution.
Journal of Mathematical Imaging and Vision.
(to appear).

Quéau, Y., Durou, J.-D., and Aujol, J.-F. (2017c).
Normal Integration: A Survey.
Journal of Mathematical Imaging and Vision.
(to appear).

Quéau, Y., Durou, J.-D., and Aujol, J.-F. (2017d).
Variational Methods for Normal Integration.
Journal of Mathematical Imaging and Vision.
(to appear).

Quéau, Y., Lauze, F., and Durou, J.-D. (2015).
Solving Uncalibrated Photometric Stereo using Total Variation.
Journal of Mathematical Imaging and Vision, 52(1):87–107.

Quéau, Y., Mecca, R., and Durou, J.-D. (2016a).
Unbiased Photometric Stereo for Colored Surfaces: A Variational Approach.
In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4359–4368.

Bibliography III

Quéau, Y., Mecca, R., Durou, J.-D., and Descombes, X. (2017e).
Photometric Stereo with Only Two Images: A Theoretical Study and Numerical Resolution.
Image and Vision Computing, 57:175–191.

Quéau, Y., Mélou, J., Castan, F., Cremers, D., and Durou, J.-D. (2017f).
A Variational Approach to Shape-from-shading Under Natural Illumination.
In *Energy Minimization Methods for Computer Vision and Pattern Recognition (EMMCVPR)*.

Quéau, Y., Modrzejewski, R., Gurdjos, P., and Durou, J.-D. (2016b).
A full photometric and geometric model for attached webcam/matte screen devices.
Signal Processing: Image Communication, 40:65–81.
Elsevier.

Quéau, Y., Pizenberg, M., Durou, J.-D., and Cremers, D. (2017g).
Microgeometry capture and RGB albedo estimation by photometric stereo without demosaicing.
In *International Conference on Quality Control by Artificial Vision (QCAV)*, volume 1338 of *Proceedings of SPIE*.

Quéau, Y., Wu, T., and Cremers, D. (2017h).
Semi-Calibrated Near-Light Photometric Stereo.
In *International Conference on Scale Space and Variational Methods in Computer Vision (SSVM)*, volume 10302 of *Lecture Notes in Computer Science*, pages 656–668.

Quéau, Y., Wu, T., Lauze, F., Durou, J.-D., and Cremers, D. (2017i).
A Non-Convex Variational Approach to Photometric Stereo under Inaccurate Lighting.
In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Woodham, R. J. (1980).
Photometric Method for Determining Surface Orientation from Multiple Images.
Optical Engineering, 19(1):139–144.