# Digital (or touch-less) fingerprint lifting using structured light

Saleh Mosaddegh[1,2], Laurent Condat[3], and Luc Brun[1,2]

[1]GREYC UMR 6072, Caen, France
[2]Normastic FR CNRS 3638, France
[3]GIPSA-lab, Grenoble, France

## Abstract

*An automatic photographic acquisition system to capture images of fingerprints on crime scenes for the public force is developed. For this, we propose a complete system, from the physical acquisition device to the software that automatically yields a 3D reconstruction of the fingerprint, i.e. the fingerprint and the surface on which the print is laid. The proposed technological solution is innovative, as it relies on a **single** captured color image of the scene, on which structured light is projected. Thus, the system is portable and as easy to use as a standard camera. Since the prototype is not finalized at the present time, our study focuses on the feasibility of the proposed solutions by means of simulations with a raytracer.*

## 1. Introduction

The analysis of fingerprints plays a major role for the police and justice agents, e.g. to establish the proof of a crime. Latent fingerprints can be invisible but are revealed by a monochromatic powder applied with a pencil. Then they are lifted using an adhesive tape, to be analyzed and identified once back in the lab. Several drawbacks are inherent to this manual method, including the possible involuntary deterioration of the fingerprint, the heavy, time-consuming and costly procedure of gathering and analyzing a large amount of fingerprints, and finally the fact that gathering the fingerprint removes it from its support, which can deprive the police of potential supplementary elements of proof.

In this work[1], we aim at proposing a 3D digital acquisition system of fingerprints, to simplify and accelerate their analysis. To do so, it is necessary to capture, simultaneously to the image of the fingerprint itself, the geometry of the 3D scene, under the form of a depth value for every pixel of the image. The resulting mesh and its texture can then be used to unwrap the fingerprint as it would have appeared on a plane, before identification. The Fig. 1 schematically represents the chal-
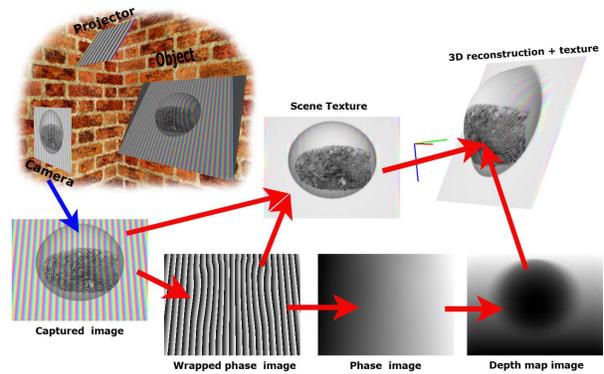


Figure 1: Schematic representation of the challenges raised by the project. The red arrows indicate the problems to solve, for which solutions are proposed in the Sections 3 and 4.

lenge raised by the project.

Numerous solutions, in the academic and industrial worlds, have been proposed, in order to capture the 3D geometry of a scene. The passive methods, such as stereoscopy, require two or more calibrated views, hence several cameras. Among the active methods, based on the detection of a specific radiation reflected by the object, some are based on laser beam scanners so that the system has to be immobile during the acquisition time of several seconds at least. We did not retain these solutions for reasons of portability and cost. Thus, the natural choice in our setting is structured light projection [9, 8], which can be implemented using cheap consumer electronics devices, like a digital camera and a LCD projector. In this method, the geometry of the scene is obtained by analyzing the deformation of the light pattern projected on the scene, as it appears in the image. The advantage of structured-light 3D scanners is speed and precision. Instead of scanning one point at a time, structured light scanners scan the entire field of view of the camera at once so they are more faster than laser scanners and also more precise, because they suffer less from the problem of distortion from motion.

There exist many different types of patterns and for each pattern, a specific reconstruction technique has been developed. Among these techniques, the projection of sinusoidal

---

fringes [7] has proved to give the best results in terms of accuracy, so that we adopted this approach for our purpose.

There are three main differences between the current techniques in the literature based on sinusoidal fringe pattern and our proposed method: we use a tele-centric camera instead of a perspective camera; we use only one color pattern instead of three gray patterns; and more importantly, we introduce an algorithm which cleans the pattern from the input image in order to recover the scene's texture (see Section 4).

We employ a tele-centric camera for two main reasons. First, using a tele-centric camera results in a sharper image of the scene's texture (fingerprint) since the depth-of-field of a tele-centric camera is very long (theoretically infinite). Second, for the acquisition system to work, we need first to calibrate the system in order to estimate the intrinsic parameters of both camera and projector, as well as their extrinsic parameters (e.g. their relative orientation and position). Using a tele-centric camera simplifies the calibration process by reducing the number of extrinsic parameters to be estimated (we only need to estimate the tilt angle of the projector, see Section 3.3).

We project only one RGB pattern instead of three sinusoidal pattern, reducing the number of acquisitions from three to only one, which means that the final system is small and portable. Simultaneous projection of three patterns through the R, G, B, channels of a single color pattern is not a new idea and it has been tried on monochromatic scenes [6] but we are the first to deal with a scene with color texture.

For the sake of portability and cost, we also recover the texture of the scene directly from the captured image without need for an extra acquisition.

Thus, the proposed acquisition system consists in a digital tele-centric camera on which a fringe pattern projector is mounted. From a **single** color image, shot on the fly without any known reference frame or geometric constraint between the object and the operator, the designed software is able to automatically retrieve the 3D geometry of the scene from the deformations of the projected pattern and also to separate the texture information (the fingerprint; the image of the scene) from the projected pattern.

## 2. Image formation model

The projector lens is not tele-centric, which means that in the scene, the sines appear locally with a frequency all the lower as the object is far away, as illustrated in Fig. 2 where $d$ is the distance from the projection point to the pattern. These deformations, induced by the perspective projection, will be used to infer the depth at every point of the image.

The color pattern projected on the scene consists in three sinusoidal fringes in the red (R), green (G) and blue (B) bands of the visible spectrum, varying along the horizontal direction, with same magnitude and phase shift of $\pm 2\pi/3$, forming a rainbow-like pattern. The light pattern is projected using a projector, tilted towards the bottom, which means that the rays with same color form vertical slices. Thus, the frequency of the fringes changes with the distance and every point of the 3D scene with coordinates $(x, y, z) \in \mathbb{R}^3$ is illuminated by a
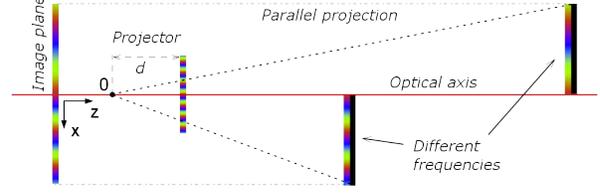


Figure 2: Schematic representation of the acquired image when the scene is made of two identical half-planes orthogonal to the optical axis. This figure illustrates the fact that the frequency of the sines depends, proportionally in this case, on the distance between the object point and the projection center of the projector.

radiance which can be written, in each of the three bands, as:

$$r^R(x,y,z) = p\big(1 + \cos(f.d.x/z - 2\pi/3)\big),$$
$$r^G(x,y,z) = p\big(1 + \cos(f.d.x/z)\big), \qquad (1)$$
$$r^B(x,y,z) = p\big(1 + \cos(f.d.x/z + 2\pi/3)\big),$$

where $p$ models the power of the light projector and $f$ is the frequency of the sines (rad/mm).

Consequently, we model each pixel value of the acquired image $\mathbf{v} = (v^R, v^G, v^B)$ by

$$v^R[\mathbf{k}] = u^R[\mathbf{k}]\big(a^R + p(1 + \cos(g[\mathbf{k}] - \tfrac{2\pi}{3}))\big),$$
$$v^G[\mathbf{k}] = u^G[\mathbf{k}]\big(a^G + p(1 + \cos(g[\mathbf{k}]))\big), \qquad (2)$$
$$v^B[\mathbf{k}] = u^B[\mathbf{k}]\big(a^B + p(1 + \cos(g[\mathbf{k}] + \tfrac{2\pi}{3}))\big),$$

where $\mathbf{k} = (k_1, k_2) \in \mathbb{Z}^2$ is the pixel location, $\mathbf{a} = (a^R, a^G, a^B)$ represents the color of the ambient light, assumed constant throughout the image, and $\mathbf{u} = (u^R, u^G, u^B)$ is the color image of the texture of the scene. This simple model assumes that the radiance, in each of the R, G, B bands, of a light ray reflected by the object, is equal to the product of the reflectance of the object at the hit point with the radiance of the incident light ray. This simplistic model of diffuse refection does not take into account the Lambertian and specular aspects of the interaction between ray lights and objects.

Thus, if the structured light projector had been switched off, the acquired image would have been

$$\tilde{v}^R[\mathbf{k}] = a^R.u^R[\mathbf{k}], \ \tilde{v}^G[\mathbf{k}] = a^G.u^G[\mathbf{k}], \ \tilde{v}^B[\mathbf{k}] = a^B.u^B[\mathbf{k}].$$

The main challenge of the project is then to recover, only from the acquired image $\mathbf{v}$, the texture image $\mathbf{u}$ and the phase image $g$, the values $a^R$, $a^G$, $a^B$ and $p$ being unknown. This is clearly an ill-posed inverse problem, all the more since in reality the image $\mathbf{v}$ is corrupted by distortions, like noise or higher harmonics to the sines due to the nonlinearities of the projector and sensor.

We note that the considered problem of recovering $g$ and $\mathbf{u}$ from $\mathbf{v}$ has common points with the classical problem of *demosaicking* in image processing [1], since the missing information (the depth map) is encoded in some way in the spectrum of the available image. This parallel motivated the choice of the color sines as *carrier waves* modulating the missing information, like in demosaicking where the chrominance is modulated by sines [1].

## 3. Analysis of the pattern deformations

### 3.1. Extraction of the wrapped phase

The first problem to tackle is the extraction of the wrapped phase at every pixel location $\mathbf{k}$; that is, the value $g[\mathbf{k}]$ modulo $2\pi$, within $]-\pi, \pi]$, for which several classical approaches exist in the literature including *Phase Stepping* [6], *Windowed Fourier transform* [10] and *Wavelet transform* [5]. Phase Stepping is valid only in the case the objects of the scene are white, or monochromatic with known color. In our case, the texture can be arbitrary. Windowed Fourier transform method is suitable in our setting, but it is sensitive to the choice of the window size, for which there is no criterion of choice.

The wavelet transform is a multiresolution tool which has shown to be very effective to analyze fringes and yields better results than the Fourier transform [5]. The Morlet wavelet is known to provide the best compromise between spatial and frequential localizations [2]. This 1D wavelet can be written as:

$$\psi(x) = \pi^{-\frac{1}{4}} e^{icx} e^{-\frac{x^2}{2}}, \qquad (3)$$

where $i^2 = -1$ and $c$ is a parameter bigger than 5. By taking the inverse tangent of the ratio between the imaginary and real parts of the wavelet transform of the image, we obtain the phase. We adopted this approach in our project, because it is robust to the distortions like the presence of noise or higher harmonics [5]. Using a 2D wavelet transform would further improve the robustness of the phase detection, at the price of a much higher computational cost [5]; we did not retain this solution because the fringes that we analyze are mostly vertical.

The procedure to extract the wrapped phase is as follows. We define a complex-valued image whose real and imaginary parts are the two chrominance channels (more details in Section 4). The wavelet analysis is performed on the rows of the image, considered as 1D signals. For every pixel, we compute the complex wavelet coefficient as the Hermitian product of the scaled wavelet centered at the pixel and the row containing the pixel. We run through a range of scales, refining the search by dichotomy to find the scale maximizing the modulus of the wavelet coefficient. At the found scale, the complex argument of the wavelet coefficient gives the phase.

### 3.2. Phase unwrapping

The phase extraction methods presented in the previous section provide wrapped phase values, between $-\pi$ and $\pi$. Thus, the wrapped phase image contains lines of discontinuities, under the form of horizontal jumps of $2\pi$. In order to recover the phase $g$, it is necessary to apply a *phase unwrapping* algorithm [3]. In our context, we suppose that the fingerprint lies on a surface without discontinuity; that is, the fingerprint is not on a stairstep. After having implemented and compared several methods, we adopted the classical method of Goldstein et al. [4], which is simple, fast, and robust.

### 3.3. Depth map computation

The last step of the analysis of the deformed fringes is the conversion of the unwrapped phase image to a depth map, in order to associate a depth value to every pixel of the extracted fingerprint image (see Section 4). We denote by $W$ and $H$ the horizontal and the vertical size of the sensor and we assume that each pixel is a square with the size of $ps$ millimeter. Suppose that, in the projector, the projector tilt angle is denoted by $\alpha$. Then, $\forall \, \mathbf{k} = (k_1, k_2) \in [0, \ldots, W-1] \times [0, \ldots, H-1]$, the pixel value $v[\mathbf{k}]$ corresponds to the point of the surface having real coordinates $(X[\mathbf{k}], Y[\mathbf{k}], Z[\mathbf{k}])$, hit by the ray parallel to the optical axis of the tele-centric camera and passing through the photosite of index $\mathbf{k}$ of the sensor. Since the camera is tele-centric, we have:

$$X[\mathbf{k}] = (k_1 - W/2)ps, \; Y[\mathbf{k}] = (k_2 - H/2)ps. \qquad (4)$$

Since $g[\mathbf{k}] = fX[\mathbf{k}]$ and we already reconstructed $g$, based on the similarity between right triangles (see Fig. 2) and taking into account the angle $\alpha$, we have:

$$Z[\mathbf{k}] = \frac{(2k_2 - h)g[\mathbf{k}]\sin(\alpha) + f(2k_1 - w)d}{2g[\mathbf{k}]\cos(\alpha)} ps. \qquad (5)$$

We note that for $k_1 = W/2$ or for the pixels which are very close to the central column of the image, $g[\mathbf{k}]$ tends to be very small and therefore, the informations $Z$ is lost. Thus, we use linear interpolation from the adjacent columns to calculate the central column of the depth map.

## 4. Extraction of the texture

Separating the texture $\mathbf{u}$, which includes the fingerprint, from the color fringes in the image $\mathbf{v}$ is an essential and critical part of the project. If this separation fails, the image of the fingerprint can not be exploited by the identification software, because its information is corrupted by the fringes. It is known that in natural images, the R, G, B channels are highly correlated. That is why for many image processing and computer vision tasks, it is preferable to work within a luminance/chrominance representation of colors instead. In this work, we adopt the orthonormal basis made by luminance, red-green chrominance and yellow-blue chrominance, which is an approximation of the way the human visual system treats the color information, according to the well-known theory of opponent colors. In this basis, the information of color natural images can be considered as decorrelated, in first approximation. More precisely, a vector $\mathbf{q}$ with components $(q^R, q^G, q^B)$ in the R, G, B basis has following components in this luminance/chrominance basis:

$$q^L = \frac{1}{\sqrt{3}}(q^R + q^G + q^B), \; q^{C_1} = \frac{1}{\sqrt{2}}(q^G - q^R),$$
$$q^{C_2} = \frac{1}{\sqrt{6}}(2q^B - q^R - q^G). \qquad (6)$$

The variational approach we propose for extracting the texture consists in seeking a texture image $\mathbf{u}$ and an ambient light vector $\mathbf{a}$ that minimize a quadratic regularized least-squares criterion. The chosen Tikhonov regularizer is separable in the basis $L, C_1, C_2$. Thus, the optimization problem we propose to solve is the following:

$$\min_{\mathbf{u}, \mathbf{a}} \mathcal{C}(\mathbf{u}, \mathbf{a}) = \sum_{[k_1, k_2] \in \Omega} |v^R[\mathbf{k}] - u^R[\mathbf{k}]b^R[\mathbf{k}]|^2 +$$
$$|v^G[\mathbf{k}] - u^G[\mathbf{k}]b^G[\mathbf{k}]|^2 + |v^B[\mathbf{k}] - u^B[\mathbf{k}]b^B[\mathbf{k}]|^2 + \qquad (7)$$
$$\lambda \|\nabla u^L[\mathbf{k}]\|^2 + \mu \|\nabla u^{C_1}[\mathbf{k}]\|^2 + \mu \|\nabla u^{C_2}[\mathbf{k}]\|^2,$$

$$(a^R, a^G, a^B) \qquad (.17, .18, .17) \qquad (.17, .18, .18) \qquad (.18, .19, .19)$$

Figure 3: Results of texture extraction and ambient light estimation after 0, 50 and 100 iterations and at convergence ($\approx 900$ iterations), from the left to the right, respectively. The true value of the ambient light is $\mathbf{a} = (0.2, 0.2, 0.2)$.

where the vector $\mathbf{b}[\mathbf{k}] = (b^R[\mathbf{k}], b^G[\mathbf{k}], b^B[\mathbf{k}])$ is defined by:

$$
\begin{aligned}
b^R[\mathbf{k}] &= \big(a^R + p(1 + \cos(g[\mathbf{k}] - 2\pi/3))\big), \\
b^G[\mathbf{k}] &= \big(a^G + p(1 + \cos(g[\mathbf{k}]))\big), \\
b^B[\mathbf{k}] &= \big(a^B + p(1 + \cos(g[\mathbf{k}] + 2\pi/3))\big),
\end{aligned}
\tag{8}
$$

and $\|\nabla u\|^2$ denotes the scalar product $< u, u * r >$ calculated with the discrete Laplacian filter

$$
r = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}.
$$

The parameters $\lambda$ and $\mu$ are important, because they control the tradeoff between the fit to the data, the smoothness of the luminance and the smoothness of the chrominance. Currently, these values are chosen empirically, depending on the noise level, with $\mu/\lambda \approx 10$ as a rule of thumb.

The solution to this quadratic optimization problem is obtained by setting to zero the partial derivatives of $\mathcal{C}$ with respect to the unknowns $u^R[\mathbf{k}]$, $u^G[\mathbf{k}]$, $u^B[\mathbf{k}]$, $a^R$, $a^G$, $a^B$. Thus, we obtain the following linear system to solve for $\mathbf{u}$:

$$
\begin{aligned}
v^R[\mathbf{k}].b^R[\mathbf{k}] = u^R[\mathbf{k}](b^R[\mathbf{k}])^2 + (\frac{\lambda}{3} + \frac{2\mu}{3})(u^R * r)[\mathbf{k}] + \\
(\frac{\lambda}{3} - \frac{\mu}{3})(u^G * r)[\mathbf{k}] + (\frac{\lambda}{3} - \frac{\mu}{3})(u^B * r)[\mathbf{k}], \\
v^G[\mathbf{k}].b^G[\mathbf{k}] = u^G[\mathbf{k}](b^G[\mathbf{k}])^2 + (\frac{\lambda}{3} - \frac{\mu}{3})(u^R * r)[\mathbf{k}] + \\
(\frac{\lambda}{3} + \frac{2\mu}{3})(u^G * r)[\mathbf{k}] + (\frac{\lambda}{3} - \frac{\mu}{3})(u^B * r)[\mathbf{k}], \\
v^B[\mathbf{k}].b^B[\mathbf{k}] = u^B[\mathbf{k}](b^B[\mathbf{k}])^2 + (\frac{\lambda}{3} - \frac{\mu}{3})(u^R * r)[\mathbf{k}] + \\
(\frac{\lambda}{3} - \frac{\mu}{3})(u^G * r)[\mathbf{k}] + (\frac{\lambda}{3} + \frac{2\mu}{3})(u^B * r)[\mathbf{k}],
\end{aligned}
\tag{9}
$$

and the linear system to solve for $\mathbf{a}$ is:

$$
\begin{aligned}
\sum_{\mathbf{k}} u^R[\mathbf{k}]v^R[\mathbf{k}] &= \sum_{\mathbf{k}} u^R[\mathbf{k}]^2(a^R + s^R[\mathbf{k}]), \\
\sum_{\mathbf{k}} u^G[\mathbf{k}]v^G[\mathbf{k}] &= \sum_{\mathbf{k}} u^G[\mathbf{k}]^2(a^G + s^G[\mathbf{k}]), \\
\sum_{\mathbf{k}} u^B[\mathbf{k}]v^B[\mathbf{k}] &= \sum_{\mathbf{k}} u^B[\mathbf{k}]^2(a^B + s^B[\mathbf{k}]),
\end{aligned}
\tag{10}
$$

where we introduced the vector $\mathbf{s}[\mathbf{k}] = (s^R[\mathbf{k}], s^G[\mathbf{k}], s^B[\mathbf{k}])$ defined by:

$$
s^R[\mathbf{k}] = p\big(1 + \cos(g[\mathbf{k}] - \frac{2\pi}{3})\big), \quad s^G[\mathbf{k}] = p\big(1 + \cos(g[\mathbf{k}])\big),
$$

$$
s^B[\mathbf{k}] = p\big(1 + \cos(g[\mathbf{k}] + \frac{2\pi}{3})\big).
\tag{11}
$$
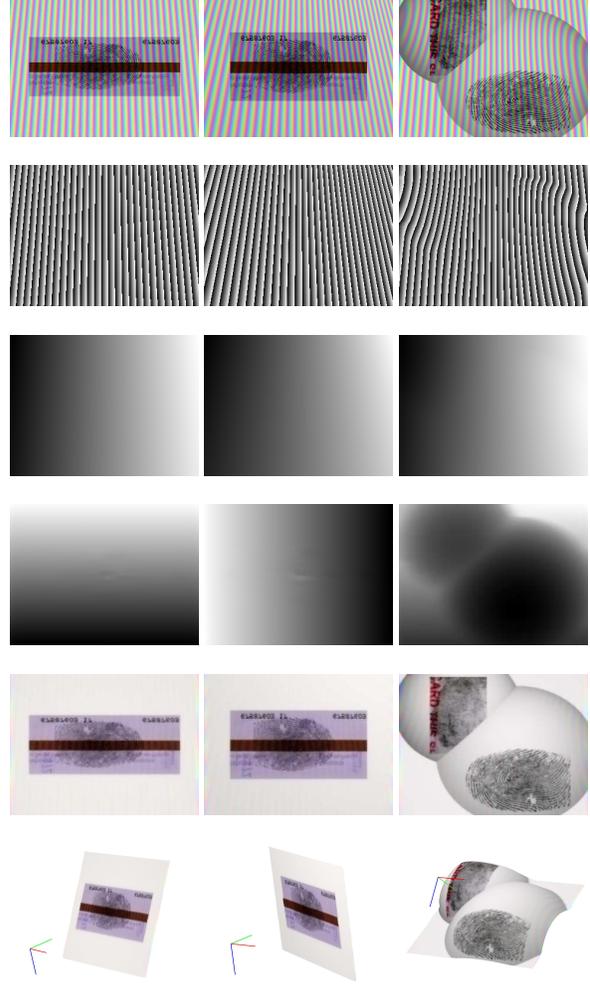


Figure 4: Output of the proposed algorithm for 3 different scenes (see the text for the explanation).

Since the optimization problem is convex, we use an alternate minimization strategy. For a given estimate of $\mathbf{a}$, we run a fixed number of iterations of the conjugate gradient method to approximately solve the linear system and find an estimate of $\mathbf{u}$. From this $\mathbf{u}$, we update $\mathbf{a}$ by a closed-form least-squares regression on the image. We iterate this process many times. The Fig. 3 shows the results of this joint estimation of $\mathbf{u}$ and $\mathbf{a}$ on a simulated image (see next section) after different numbers of total iterations of the conjugate gradient. We notice that the process converges, which can be proved theoretically from the convexity of the problem. The convergence speed highly depends on $\mu$ and $\lambda$ and remains to be studied more precisely. Note that, as $p$ is unknown, we actually recover $\mathbf{a}$ and $\mathbf{u}$ up to a constant, which does not play any role.

## 5. Simulations

In order to test the different steps of the analysis process, we implemented a simulator based on a rendering engine of 3D scenes by raytracing. This simulator allows to take into con-

sideration the color and nature of the surface material to mimic the physical reflection and refraction of the light rays. The images in Fig. 1 are the results of the algorithms described in this article, where the surface is the compound of a sphere and a plane orthogonal to the optical axis. Also, each column in the Fig. 4 shows the obtained results from simulations with different scenes; a plane tilted along the horizontal axis; a plane tilted along the vertical axis and a scene composed of two spheres cut by a plane tilted along the horizontal axis. The rows are, from top to bottom, corresponding to the simulated images $\mathbf{v}$, the wrapped phase images (where black represents the value $-\pi$ and white the value $\pi$), the phase images, the depth map images, the extracted texture images and finally the 3D reconstructions of the scenes. We simply applied a median filter on the Z coordinates of the point cloud to smooth the surface. Using advanced surface fitting techniques results in smoother surfaces.

The source code is freely available to download from the website of the first author[2].

## 6. Conclusion

In this paper we proposed a complete system for image acquisition and 3D reconstruction of fingerprints, using a digital camera and a projector of structured light with a sinusoidal fringe pattern. The main contributions of our proposed technique are threefold: we use a tele-centric camera, simplifying the calibration and reconstruction processes; we project only one color pattern on the textured scene, increasing the speed and portability of the system and decreasing its cost, and finally we introduce a variational approach and use the same input image in order to recover the scene's texture.

The variational approach developed to jointly extract the texture and estimate the ambient light has shown its efficiency and robustness. Ongoing tests of performance should make the choice of the model parameters more precise. Alternative solutions based on the paradigm of sparsity in signal processing or on total variation minimization should be studied.

The simulation results show that the system design and the developed image processing methods are viable and relevant. The project sheds light on a promising digital alternative to the traditional manual lift of fingerprints.

The potential applications of our methodology are numerous and go far beyond the lift of fingerprints. Arbitrary 3D objects could be captured this way, e.g. 3D human faces for biometry or entertainment.

## References

[1] L. Condat. A new color filter array with optimal properties for noiseless and noisy color image acquisition. *IEEE Transactions on Image Processing*, 20(8):2200–2210, 2011. 2

[2] A. Dursun, S. Ozder, and F. N. Ecevit. Continuous wavelet transform analysis of projected fringe patterns. *Measure Science and Technology*, 2004. 3

[3] D. C. Ghiglia and M. D. Pritt. *Two-Dimensional Phase Unwrapping: Theory, Algorithms and Software*. A Wiley-Interscience Publication, 1998. 3

[4] R. M. Goldstein, H. A. Zebker, and C. L. Werner. Satellite radar interferometry: Two-dimensional phase unwrapping. *Radio science*, 23(4), 1988. 3

[5] L. Huang, Q. Kemao, B. Pan, and A. K. Asundi. Comparison of Fourier transform, windowed Fourier transform, and wavelet transform methods for phase extraction from a single fringe pattern in fringe projection profilometry. *Optics and Lasers in Engineering*, 48:141–148, 2010. 3

[6] P. S. Huang, Q. Hu, F. Jin, and F.-P. Chiang. Color-encoded digital fringe projection technique for high-speed three-dimensional surface contouring. *Optical Engineering*, 38(6):1065–1071, 1999. 2, 3

[7] N. Karpinsky and S. Zhang. High-resolution, real-time 3d imaging with fringe analysis. *J. Real-Time Image Process.*, 7(1):55–66, Mar. 2012. 2

[8] J. Salvi, S. Fernandez, T. Pribanic, and X. Llado. A state of the art in structured light patterns for surface profilometry. *Pattern Recognition*, 43, 2010. 1

[9] J. Salvi, J. Pagès, and J. Batlle. Pattern codification strategies in structured light systems. *Pattern Recognition*, 37, 2004. 1

[10] M. Takeda, H. Ina, and S. Kobayashi. Fourier-transform method of fringe pattern analysis for computer-based topography and interferometry. *Journal of the Optical Society of America*, 72, 1982. 3

---

[2]http://salehas.perso.sfr.fr/structuredLight.html