Variational approach to moiré pattern synthesis

Guy Lebanon

*Language Technologies Institute, School of Computer Science, Carnegie Mellon University, 5000 Forbes Avenue,
Pittsburgh, Pennsylvania 15213*

Alfred M. Bruckstein

Department of Computer Science, Technion–Israel Institute of Technology, Haifa 32000 Israel

Moiré phenomena occur when two or more images are nonlinearly combined to create a new superposition image. Moiré patterns are patterns that do not exist in any of the original images but appear in the superposition image, for example as the result of a multiplicative superposition rule. The topic of moiré pattern synthesis deals with creating images that when superimposed will reveal certain desired moiré patterns. Conditions that ensure that a desired moiré pattern will be present in the superposition of two images are known; however, they do not specify these images uniquely. The freedom in choosing the superimposed images can be exploited to produce various degrees of visibility and ensure desired properties. Performance criteria for the images that measure when one superposition is better than another are introduced. These criteria are based on the visibility of the moiré patterns to the human visual system and on the digitization that takes place when the images are presented on discrete displays. We propose to resolve the freedom in moiré synthesis by choosing the images that optimize the chosen criteria. © 2001 Optical Society of America

OCIS codes: 120.4120, 100.2000, 100.2650, 100.3190.

1. INTRODUCTION

The term moiré comes from French, where it refers to watered silk. The moiré silk consists of two layers of fabric pressed together. As the silk bends and folds, the two layers shift with respect to each other, causing the appearance of interfering patterns. The moiré technique for manufacturing cloth was developed in China a long time ago and was later introduced to France in 1754 by the English manufacturer Badger.[1] Natural moiré phenomena can be seen in daily life, for example in the folds of a moving nylon curtain or in looking through parallel wire-mesh fences. The first scientific observations were made by Lord Rayleigh,[2] who suggested using the moiré phenomenon for testing the quality of gratings.

Two goals exist in moiré pattern research. The first is the analysis of moiré patterns. This usually involves some physical situation in which moiré patterns appear either naturally or by human intervention. The task is to analyze and characterize the patterns. Most of the research in moiré pattern analysis deals with finding equations that describe the moiré patterns. In moiré pattern synthesis the generation of certain moiré patterns is required. The synthesis process involves producing two images such that when these images are superimposed the required moiré patterns emerge. Moiré synthesis and analysis are tightly linked, and understanding one task gives insight into the other.

Over the years, various methods to model and analyze the moiré phenomenon have been suggested. In Section 2 we describe two main approaches to modeling the moiré phenomenon. In Section 3 we present the moiré pattern synthesis problem, and in Section 4, we introduce criteria for measuring the performance of moiré patterns in a superposition. In Section 5 we discuss the integrability constraint that ensures that a certain vector field is a gradient field. Section 6 reviews some basic results from variational calculus and their use in moiré synthesis, and Section 7 addresses the problem of recovering the potential function of a gradient field. Section 8 contains some experimental results, followed by conclusions in Section 9.

Throughout this paper, for the sake of simplicity we discuss the case of superposition of two images. It is not too difficult to extend the results to several superimposed images. We assume a multiplicative superposition rule. Such a rule is motivated by the multiplication effect implicit in laying transparencies on one another. It is possible, however, to consider other superposition rules.[3] The nonlinearity of the multiplicative superposition allows new frequencies that do not exist in the original image to appear in the superimposed image. In fact, nonlinearity is at the heart of the moiré phenomenon, and linear superposition such as addition does not elicit it (see Fig. 1).

2. MOIRÉ PATTERN ANALYSIS

Two models for moiré pattern analysis are reviewed. The indicial equation method operates in the image plane, and the Fourier domain method operates in the frequency plane. Detailed descriptions of these models appear in Ref. 4.

A. Indicial Equation Method

The simplest and oldest model for analyzing the geometric shape of moiré patterns in the superposition of two curvilinear gratings is the indicial (or parametric) equa-
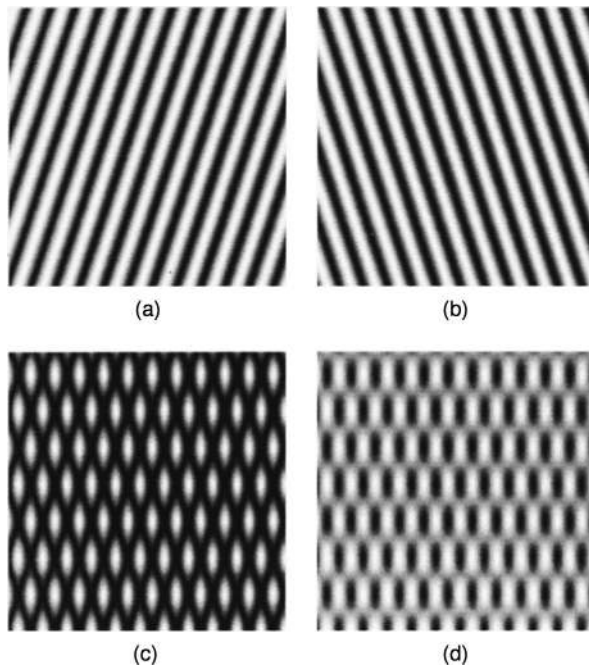
Fig. 1. Two types of superposition. The two gratings in (a) and (b) are multiplied in (c) and added in (d). Image brightness is scaled.

tion method surveyed in Refs. 1 and 5. This model is based on the curve equations of the original curvilinear gratings. If each of the original layers is regarded as an indexed family of curves, the moiré pattern of the superposition forms a new indexed family of curves, whose equations can be inferred from the equations of the original gratings.

According to this model the original images consist of black curves on a white background. The curves in each image are assumed to be the equal-height contours of two-dimensional (2D) functions. We thus have two images that consist of black curved gratings whose center lines are the equal-height contours of two functions $\psi(x, y)$ and $\phi(x, y)$ as follows:

$$\psi(x, y) = m, \qquad m \in \mathbb{Z},$$

$$\phi(x, y) = n, \qquad n \in \mathbb{Z}. \qquad (1)$$

Since the images are binary, the multiplicative superposition is also an AND operator. Each of the curves in both images has an index given by the height of the respective contour. As a result, adjacent curves get adjacent integers as indices. We denote the indices of the curves on one image by an integer variable m and the indices of the curves in the second image by the integer variable n. The coordinates m,n define an (m,n) net. At each point of the (m,n) net, an m and an n curve intersect.

The (k_1, k_2) moiré curves are defined as the curves joining the intersection points of m and n curves whose indices obey $k_1 m + k_2 n = l$, where k_1, k_2 are constant integers and l runs over the set of integers.

Conceptually, if we let m and n vary continuously, the (k_1, k_2) moiré curves that obey

$$k_1 m + k_2 n = l, \qquad l \in \mathbb{Z}, \qquad (2)$$

become continuous curves that may be regarded as equal-height contours of a new bivariate function $g(x, y)$.

The order of the (k_1, k_2) moiré is defined to be the highest absolute value of k_1, k_2. The first-order moirés are therefore the additive moiré $u + v = l$ and the subtractive moiré $u - v = l$. In other words, first-order moiré patterns are the curves connecting the intersection points of the constant sum and the constant difference of the curves' index values.

Clearly, not all the (k_1, k_2) moirés stand out visually. The visibility of the moiré patterns will be discussed at greater length below. Usually only first-order moirés stand out, if at all. In the case of first-order moirés, sometimes only the additive or only the subtractive moiré stands out, and sometimes no moiré is apparent at all. (For an intuitive explanation of the visibility of the first-order moirés see Ref. 4, p. 1.)

Substituting Eqs. (1) into Eq. (2) results in eliminating the indices m,n:

$$k_1 \psi(x, y) + k_2 \phi(x, y) = l, \qquad l \in \mathbb{Z}. \qquad (3)$$

We can thus state that the center lines of the (k_1, k_2) moiré corresponds to the equal-height contours of $g(x, y)$ $= k_1 \psi(x, y) + k_2 \phi(x, y)$. The indicial equation method permits a complete geometric specification of the moiré curves based on the implicit geometric specification of the curves in the two original images as equal-height contours of bivariate functions.

The indicial equation method has several drawbacks. In order to use the indicial equation, we need explicit analytic expressions for the curves, which may not be readily available. Although the method gives us a condition for the possible moiré curves, it does not tell us whether the moiré pattern will indeed be visible to the human eye. For example, consider the superposition of a family of horizontal lines and a family of vertical lines. The superposition will consist of small squares. There will not be any visible pattern, although the first-order moiré patterns are diagonal lines connecting the opposite vertices of each square (see Fig. 2).

B. Fourier Domain Method

So far, we have assumed that the original images were binary images showing the equal-height contours of $\psi(x, y)$
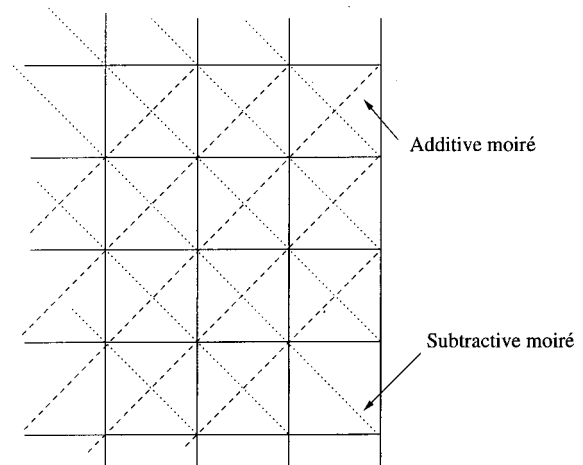

Fig. 2. Nonvisible additive and subtractive moiré patterns.

and $\phi(x, y)$ as black curves on a white background. We now generalize this setting through the concept of periodic profile.

The original images are allowed to be $p(\psi(x, y))$ and $p(\phi(x, y))$, where $p(z)$ is a periodic function of one variable with period of 1. If $p(z)$ is taken to be a discrete impulse train [$p(z) = 1$ for $z \in \mathbb{Z}$ and 0 otherwise] the images $p(\psi(x, y))$ and $p(\phi(x, y))$ reduce to binary images with curves representing equal-height contours of ψ and ϕ.

In this section we will restrict ourselves to linear ψ and ϕ: $\psi(x, y) = p_1 x + q_1 y$, $\phi(x, y) = p_2 x + q_2 y$, p_1, p_2, $q_1, q_2 \in \mathbb{R}$. Note that $p(\psi(x, y))$ and $p(\phi(x, y))$ will be periodic images. In Fig. 3(a) a linear function $\psi(x, y)$ is shown. In Figs. 3(b), 3(c), and 3(d), $p(\psi(x, y))$ is shown, where the periodic profile p is a discrete-impulse-train, a cosine, and a square-wave grating, respectively. Although we restrict ourselves here to linear functions, results obtained in this case are useful, since many functions can locally be approximated by a linear function (the first two terms of the 2D Taylor expansion).

The Fourier domain method analyzes the moiré pattern in the frequency domain. According to the convolution theorem, the multiplicative superposition rule in the image domain transforms to a two-dimensional convolution between the spectra of the original images:

$$g(x, y) = p(\psi(x, y))p(\phi(x, y)) \Leftrightarrow G(u,v)$$
$$= \mathcal{F}[p(\psi(x, y))] * \mathcal{F}[p(\phi(x, y))], \quad (4)$$

where the asterisk denotes convolution.

The Fourier transform of $p(x) = \exp(2\pi j f x)$ is $P(u) = \delta(u - f)$. In the 2D domain the Fourier transform of $p(x, y) = \exp[2\pi j(f_1 x + f_2 y)]$ is $P(u,v) = \delta(u - f_1, v - f_2)$. Since p is a periodic function, it can be expanded to a Fourier series:

$$p(x, y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_{m,n} \exp[2\pi j(m u_0 x + n v_0 y)].$$

The Fourier transform of $p(x, y)$ is readily obtained from the Fourier series decomposition and linearity of the Fourier transform:

$$P(u, v) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_{m,n} \delta(u - m u_0, v - n v_0).$$

Furthermore, by the convolution theorem, $G(u,v)$ consists of translated and scaled impulses, as well.

Since $\psi(x, y) = p_1 x + q_1 y$, and $\phi(x, y) = p_2 x + q_2 y$, it follows that $p(\psi(x, y))$ and $p(\phi(x, y))$ have frequency components only in the direction of the gradients $\nabla \psi$ and $\nabla \phi$. This means that the frequency domain representation of $p(\psi(x, y))$ and $p(\phi(x, y))$ will have impulses only along the lines $p_1 u + q_1 v = 0$ and $p_2 u + q_2 v = 0$. This can also be seen from the fact that $p(\psi)$ and $p(\phi)$ are rotated one dimensional (1D) periodic functions on the x axis. Therefore $\mathcal{F}[p(\psi(x, y))]$ and
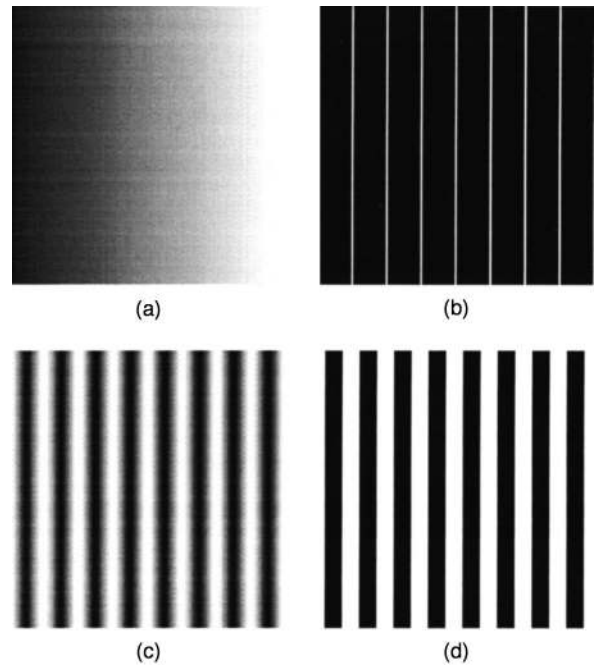

Fig. 3.   Periodic profiles of a linear function.

$\mathcal{F}[p(\phi(x, y))]$ are obtained by rotating the 1D spectra from the u axis by the same angles. In the case of a raised cosine profile, $p(\cdot)$ is given by $p(\psi(x, y)) = \frac{1}{2}\cos[2\pi\psi(x, y)] + \frac{1}{2}$, and only three impulses will exist in the frequency domain. Two impulses at either side of the origin are contributed by the cosine function. These two impulses are at distance 1 from the origin and lie along the line $p_1 u + q_1 v$, where $\phi = p_1 x + q_1 y$. The third impulse is contributed by the constant term and lies at the origin.

The convolution of the impulse spectra is performed as a discrete convolution. The location of each impulse in the superposition spectrum will be the vectorial sum of the locations of two impulses, one from each original image. We will label the (k_1, k_2) superposition impulse as the impulse whose location is created by the vectorial sum of the k_1 impulse in the first original spectrum and the k_2 impulse in the second original spectrum. The amplitude of the (k_1, k_2) impulse is the product of the amplitudes of the k_1 impulse in the first original spectrum and the k_2 impulse in the second original spectrum. Figure 4 shows the superposition of two linear gratings with a raised cosine profile in the frequency domain.

Each impulse in the 2D spectrum is characterized by three main properties: its label, its geometric location, and its amplitude. To the geometric location of an impulse, a frequency vector **f** is attached (in the frequency domain). This vector can be expressed in polar coordinates (f, θ) where f is the distance of **f** from the origin and θ is the angle of **f**. In terms of the image domain, the geometric location of an impulse in the spectrum determines the frequency f and the direction θ of the corresponding periodic component in the image. The amplitude of the impulse represents the intensity of that periodic component in the image.

The geometric locations that include impulses in the superposition spectrum but do not include impulses in

the original spectra represent the moiré patterns. These impulses represent new frequency components created by the superposition and not by one of the original images alone. Impulses that are labeled $f_{(0, i)}$ or $f_{(i, 0)}$ for some i exist in one of the original images since f_0 represents the dc term.

In the case of other profiles such as the square-wave one, we will have also higher-order moiré impulses. The spectrum of a linear square-wave grating is an infinite set of impulses along the frequency direction. Convolution of two such impulse trains will result in an infinite lattice covering the entire frequency plane.

However, we will show that the amplitude of impulses away from the origin tends to zero. We can disregard impulses with low amplitude since their effect on the image is small.

More formally, let $p(z)$ be a periodic function whose values are bounded between 0 and 1. Then all its Fourier series coefficients (impulse amplitudes) have absolute values between 0 and 1:
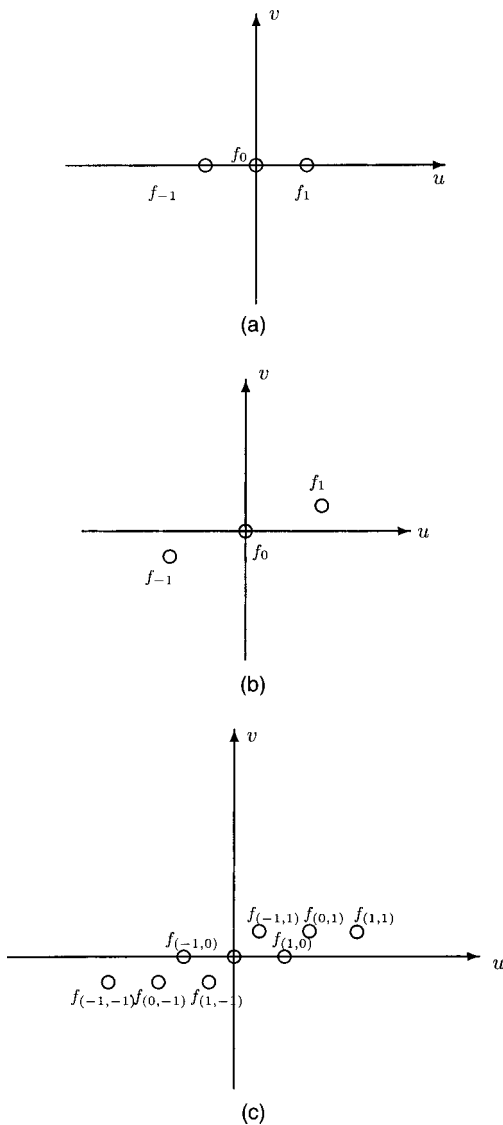
(a)

(b)

(c)

Fig. 4. Superposition in the frequency domain: spectra of the (a) first raised cosine grating, (b) second raised cosine grating, (c) superposition.

$$p(x) = \sum_{m=-\infty}^{\infty} c_m \exp(2\pi jmfx),$$

$$0 \leq |c_m| = \left| \frac{1}{T} \int_T p(x) \exp(2\pi jmfx) \mathrm{d}x \right|$$

$$\leq \left| \frac{1}{T} \int_T \exp(2\pi jmfx) \mathrm{d}x \right| \leq 1.$$

Furthermore, it is true for any convergent Fourier series that $|c_m| \to 0$ as $|m|$ tends to ∞. Moreover, if $p(x)$ has n continuous derivatives, then its Fourier transform $P(u)$ tends to 0 as $|u| \to \infty$ at least as fast as $1/u^{n+1}$ (see Ref. 6, p. 74): $\lim_{|u|\to\infty}|P(u)| = O[1/(u^{n+1})]$. Thus the smoother the function, the more rapidly the coefficients of the series tend to zero.

Recall that the spectra of the original images are given by $P(u)$ rotated along the frequency directions. Also, the amplitude of the (k_1, k_2) impulse is the product of the amplitudes of the k_1 and k_2 impulses in the original images' spectra. Therefore the amplitude of the (k_1, k_2) impulse in the superposition will tend to 0 faster than $[1/(u_1, u_2)]^{n+1}$, where u_1 and u_2 are the locations of the k_1 and k_2 impulses, respectively.

The Fourier approach has the advantage that it enables us to analyze the moiré patterns in the frequency domain, which is a suitable domain for determining the response of the visual system.

In the case of nonlinear gratings, the spectrum of the gratings no longer consists of impulses and may be continuous. It is then impossible to analyze the superposition spectrum with the same ease as before. However, a local frequency analysis can use the above results since any smooth function can be approximated by a linear function in a small enough region.

3. PROBLEM OF MOIRÉ SYNTHESIS

Synthesis of moiré patterns is the generation of two images that when superimposed will reveal the intended moiré pattern. We restate the condition for the generation of a certain moiré pattern. Let $\psi(x,y), \phi(x,y)$ be two 2D functions, and let $p(z)$ be a 1D periodic function with period of unity. The superposition $p(\psi(x,y))p(\phi(x,y))$ will contain the (k_1, k_2) moiré pattern whose geometric layout is the equal-height contours of $g(x,y)$ if the following condition holds:

$$g(x,y) = k_1\psi(x,y) + k_2\phi(x,y). \tag{5}$$

In the synthesis setting we assume that $g(x,y)$ is given. Equation (5) shows that ϕ is completely determined by g and ψ (and, conversely, ψ is completely determined by g and ϕ). There is, however, a degree of freedom in choosing ψ or ϕ. The moiré pattern described by g will stand out visually for some choices of ψ, and yet it may be hidden for other choices owing to the inherent filtering process of the human visual system. We will first determine criteria for evaluating the superposition image. Using these criteria, we will choose ψ and ϕ that will satisfy Eq. (5) and optimize our performance criteria. From here on, we will deal with synthesis of first-order $(1, -1)$ moiré

patterns, since higher-order moiré patterns are generally less dominant (see Subsection 2.B). Synthesis of the (1, 1) or higher-order moiré patterns requires some straightforward modifications.

4. PERFORMANCE CRITERIA

When synthesizing moiré patterns, we have to choose ψ and ϕ from an infinite set of functions that satisfy Eq. (5). In this section we propose criteria to estimate the visual performance of the moiré. On the basis of these criteria we can decide whether one choice of ψ and ϕ that satisfy Eq. (5) is better than another. In Section 7 we use these criteria to optimally choose ψ and ϕ. We will first consider performance criteria for linear moiré. We assume that the superposition image and the resulting moiré patterns are approximately linear in a small enough region. On the basis of local analysis and the results for linear moiré performance, we will formulate criteria for general moiré patterns.

A. Linear Moiré Performance

The filtering process of the human visual system is nicely demonstrated in two simple experiments. A figure of cosine vertical bars with continuously varying frequencies and amplitudes (see Ref. 7, Fig. 2.4-3, p. 35) demonstrates sensitivity to the spatial frequency of the cosine bars. Comparing a figure of a checkerboard with a rotated duplicate (see Ref. 8, Fig. 5.2, p. 83) demonstrates the nonisotropy of the visual system filtering: The human visual system is more sensitive at 0 and 90 deg than at 45 deg to changes with equal contrast and frequency.

The contrast of a pattern I is defined by $C = (I_{max} - I_{min})/(I_{max} + I_{min})$, where I_{max} and I_{min} are the maximum and minimum intensities, respectively, in the pattern. For an absolute uniform image, $I_{max} = I_{min}$ and $C = 0$. For a square-wave grating, $I_{max} = 1$, $I_{min} = 0$, $C = 1$. Since the denominator is proportional to the mean intensity, contrast can also be considered as the degree of modulation of intensity above the mean. For sinusoidal gratings the contrast is proportional to the amplitude.

The contrast sensitivity function (CSF) of a system is CSF=output contrast/input contrast. It is not always feasible to measure the output for humans in a completely controlled fashion. Necessarily, psychological experiments are used, requiring many assumptions about the system's behavior.

The experimental procedure for measuring the CSF involves presenting each of a set of vertical sinusoidal gratings on a visual display to a viewer who can vary the contrast control while maintaining contrast average luminance. For a given pattern coarseness, the viewer is asked to adjust the contrast until the grating is just barely distinguishable. The threshold of contrast perception $c(u)$ is obtained at different spatial frequencies, and the contrast sensitivity function is $\text{CSF}(u) = \alpha/c(u)$ (the constant α is assigned to barely distinguishable contrast).

Dooley (see Ref. 9, p. 118) has provided the following equation to fit the data from the above experiments:

$$\text{CSF}(u) = |5.05[\exp(-0.138u)][1 - \exp(0.1u)]|, \quad (6)$$

where $u = 2\pi f$ and f is the spatial frequency along the x axis in cycles per degree.

A review of the nonisotropy of the visual system[10] describes results of the following experiment. Gratings were presented to a viewer at different angles and different distances. The distances at which the gratings were barely visible represent the sensitivity of the visual system to orientation.

As a measure of the visibility of an impulse whose location on the frequency plane is $\mathbf{f}$, we take

$$V(\mathbf{f}) = H_1(\|\mathbf{f}\|) \cdot H_2(\text{angle}(\mathbf{f})), \quad (7)$$

where H_1 and H_2 are the functions obtained from the above two experiments.

Recall that for the raised cosinusoidal profile we have only the first-order moiré. In moiré pattern synthesis we receive the desired pattern $p(g(x, y)) = p(p_1 x + q_1 y)$ as input. The gradient of the desired pattern $\nabla g = (p, q)$ points in the required frequency direction. In addition, $1/\|\nabla g\|$ is the distance between two adjacent periods of $p(g(x, y))$. Therefore the magnitude of the spatial frequency of $p(g(x, y))$ is $\|\nabla g\|$. In other words, for $p(g(x, y))$ to appear as the $(1, -1)$ moiré, $\mathbf{f}_{(1, -1)}$ should be equal to ∇g. Since in moiré synthesis we receive g as input, the location of $\mathbf{f}_{(1, -1)}$ is set by ∇g when Eq. (5) is satisfied.

The freedom in choosing different ψ and ϕ that satisfy Eq. (5) allows control of the location of $\mathbf{f}_{(1, 1)}$. For the $(1, -1)$ moiré to be visible, we should minimize the visibility of the $(1, 1)$ moiré. The optimal ϕ_{opt}, ψ_{opt} for this minimization is

$$\phi_{opt} = \arg\ \min_{\phi}\ H_1(\|\mathbf{f}_{(1, 1)}\|) \cdot H_2(\text{angle}(\mathbf{f}_{(1, 1)})), \quad (8)$$

$$\psi_{opt} = g + \phi_{opt}. \quad (9)$$

$\mathbf{f}_{(1, 1)}$ can be computed by

$$\mathbf{f}_{(1, 1)} = \mathbf{f}_\phi + \mathbf{f}_\psi = \nabla\phi + \nabla\psi = 2\nabla\phi + \nabla g,$$

where we used the following result:

$$g = \psi - \phi \Rightarrow \nabla\psi = \nabla g + \nabla\phi.$$

According to the visibility function (7) the visual system will be less responsive to higher frequencies, and hence the performance will improve as $\|\nabla\phi\|$ is increased. Such uncontrolled improvement in the performance becomes problematic when we use digital media to represent the images. As we further increase $\|\nabla\phi\|$ we will get an additional strong unwanted moiré between the grating and the pixel frequency of the display, as a result of aliasing.

To account for this effect in the performance criteria, another term, M, will be added to Eq. (7) as follows:

$$V(\mathbf{f}_{(1, 1)}) = H_1(\|\mathbf{f}_{(1, 1)}\|) \cdot H_2(\text{angle}(\mathbf{f}_{(1, 1)})) + M(\|\mathbf{f}_{(1, 1)}\|). \quad (10)$$

This term will become dominant for very high frequencies and prevent the unbounded decrease in Eq. (10). This digitization term $M(\|\mathbf{f}_{(1, 1)}\|)$ should be negligible for low frequencies and dominant for high frequencies. In addition to choosing M with these properties, we should choose the crossing point

$$M(\|\mathbf{f}_{(1,1)}\|) = H_1(\|\mathbf{f}_{(1,1)}\|) \cdot \min H_2(\text{angle}(\mathbf{f}_{(1,1)}))$$
(11)

with care. To do so, we chose the function $M(\cdot)$ to be of the form $M(\cdot) = m\tilde{M}(\cdot)$, where $\tilde{M}(\cdot)$ is an increasing polynomial and m is a parameter. m is determined so as to set the crossing point (11) as described below. We define the digitization threshold T_f as the frequency at which the gratings ψ and ϕ create the aliasing effects.

We denote $\tilde{\phi}$ and $\tilde{\psi}$ as the functions computed by Eqs. (8) and (9). We would like $\|\nabla\tilde{\phi}\|$ and $\|\nabla\tilde{\psi}\|$ to be smaller than the digitization threshold by $\epsilon_1 > 0 \|\nabla\tilde{\phi}\| < T_f - \epsilon_1$, $\|\nabla\tilde{\psi}\| < T_f - \epsilon_1$. Since the choice of M affects the choice of $\tilde{\phi}$ and only then is $\tilde{\psi}$ computed, we will explore the relation between $\|\nabla\phi\|$ and $\|\nabla\psi\|$:

$$g = \psi - \phi,$$

$$\nabla\psi = \nabla g + \nabla\phi,$$
(12)

$$\|\nabla\psi\| = \|\nabla g + \nabla\phi\| \leq \|\nabla g\| + \|\nabla\phi\|.$$
(13)

If the condition $\|\nabla\tilde{\phi}\| < T_f - \epsilon_1$ holds, we have $\|\nabla\tilde{\psi}\| \leq T_f - \epsilon_1 + \|\nabla g\|$.

We denote the frequency of the crossing point as f_{cp}. If we assume that $\|\nabla\tilde{\phi}\| \leq f_{cp} + \epsilon_2$, $\epsilon_2 > 0$, we arrive at the following result: If we choose the crossing-point frequency f_{cp} according to $f_{cp} \leq T_f - \|\nabla g\| - \epsilon_1 - \epsilon_2$, $\tilde{\phi}$ and $\tilde{\psi}$ will satisfy $\|\nabla\tilde{\phi}\| \leq T_f - \epsilon_1 - \|\nabla g\|$, $\|\nabla\tilde{\psi}\| \leq T_f - \epsilon_1$.

Intuitively, the value of ϵ_1 represents how much we would like to stay away from the digitization threshold, and ϵ_2 represents the possibility that the minimization procedure will carry $\|\nabla\tilde{\phi}\|$ beyond the crossing point.

B. Performance of Nonlinear Moiré Patterns
The visibility of the (1, 1) moiré in a general superposition over a region Ω is defined as

$$W(\mathbf{f}_{(1,1)})(\Omega) = \int\int_\Omega V(\mathbf{f}_{(1,1)}(x,y))\mathrm{d}x\mathrm{d}y,$$

where $V(\mathbf{f}_{(1,1)}(x,y))$ is the function defined in Eq. (10). This follows from giving equal importance to the visibility of all points and using $V(\mathbf{f}_{(1,1)}(x,y))$ as the visibility criterion at the point (x,y).

Over a discrete image I of size $M \times N$ we have

$$W(\mathbf{f}_{(1,1)})(I) = \sum_{i=1}^{M}\sum_{j=1}^{N} V(\mathbf{f}_{(1,1)}(i,j)).$$
(14)

5. INTEGRABILITY CONSTRAINT
In minimization of the visibility of the (1, 1) moiré, Eq. (14) depends on $\mathbf{f}_{(1,1)}(x,y) = 2\nabla\phi(x,y) + \nabla g(x,y)$. Since Eq. (14) depends explicitly on $\nabla\phi$, we may state the optimization problem as follows: for each (i,j) find $\nabla\phi_{\mathrm{opt}}(i,j)$ that will minimize $V(\mathbf{f}_{(1,1)}(i,j))$.

The problem with such a scheme is that the obtained vector field $\nabla\phi_{\mathrm{opt}}$ may not be a conservative field. This means that no function can be found for which $\nabla\phi_{\mathrm{opt}}$ will be the gradient (for more information on conservative

fields and the integrability constraint see any standard vector analysis textbook such as Ref. 11, pp. 1076–1082).

Enforcing the integrability test for $\nabla\phi$, we are led to the following problem:

$$\nabla\phi_{\mathrm{opt}} = \arg\min_{\nabla\phi} W(\mathbf{f}_{(1,1)})(I)$$
(15)

$$\text{subject to} \quad \phi_{xy} = \phi_{yx}.$$

As we shall see in Section 6, our variational solution will not impose integrability as a hard constraint, but we shall enforce it approximately by means of a penalty term.

6. RESULTS FROM VARIATIONAL CALCULUS
In this section we will state some results from variational calculus that are used in the following sections. For a more complete description refer, for example, to Ref. 6 or Ref. 12.

The calculus of variations deals with minimizing functionals. A functional is a mapping from a set of functions to the real line. A fundamental result of the calculus of variations is that the extrema of functionals must satisfy an associated differential equation called the Euler equation over the domain.

The Euler equation is a necessary but not sufficient condition for the existence of an extremum. By extrema we mean local minima, maxima, and inflection points. We assume that all functions and functionals are continuous and have derivatives. Another assumption is that the functional values are positive.

For example, a functional I_1 that depends on a bivariate function $z(x,y)$, as follows,

$$I_1[z] = \int\int_\Omega F(x,y,z,z_x,z_y)\mathrm{d}x\mathrm{d}y,$$

yields the following Euler equation:

$$F_z - \frac{\partial}{\partial x}F_{z_x} - \frac{\partial}{\partial y}F_{z_y} = 0.$$

A different functional I_2 that depends on $z(x,y)$ and its derivatives, as follows,

$$I_2[z] = \int\int_\Omega F(x,y,z,z_x,z_y,z_{xx},z_{xy},z_{yy})\mathrm{d}x\mathrm{d}y,$$

yields the following Euler equation:

$$F_z - \frac{\partial}{\partial x}F_{z_x} - \frac{\partial}{\partial y}F_{z_y} + \frac{\partial^2}{\partial x^2}F_{z_{xx}}$$

$$+ \frac{\partial^2}{\partial x\partial y}F_{z_{xy}} + \frac{\partial^2}{\partial y^2}F_{z_{yy}} = 0.$$

The functional I_3 that depends on the gradient of $z(x,y)$, i.e., $\nabla z(x,y) = (p(x,y), q(x,y))$, as follows,

$$I_3[p,q] = \int\int_\Omega F(x,y,p,q,p_x,p_y,q_x,q_y)\mathrm{d}x\mathrm{d}y,$$

yields a coupled set of differential Euler equations:

$$F_p - \frac{\partial}{\partial x} F_{p_x} - \frac{\partial}{\partial y} F_{p_y} = 0,$$

$$F_q - \frac{\partial}{\partial x} F_{q_x} - \frac{\partial}{\partial y} F_{q_y} = 0.$$

The Euler differential equations need boundary conditions in order to have a specified solution. However, in many problems there are no imposed prior conditions on the boundary values, or the behavior of the function at the boundary may be restricted by some general conditions. In such cases, the variational calculus supplies us with further conditions for the boundary values. These conditions are also necessary conditions for the functional to be stationary with respect to variations (see Ref. 6, p. 208). Such conditions are called natural boundary conditions.

In the case of I_1, the natural boundary condition is

$$(F_{z_x}, F_{z_y}) \cdot \mathbf{n} = 0,$$

where $\mathbf{n}$ is the normal to the parametric curve representing the boundary of Ω.

For I_3, the natural boundary conditions are

$$(F_{p_x}, F_{p_y}) \cdot \mathbf{n} = 0, \qquad (F_{q_x}, F_{q_y}) \cdot \mathbf{n} = 0. \qquad (16)$$

Recall [from Eq. (10)] that the visibility of the (1, 1) moiré in a small area surrounding (x, y) is expressed by $V(p(x, y), q(x, y))$, where $(p(x, y), q(x, y)) = \nabla \phi$. Instead of minimizing a term based on V subject to the integrability constraints $p_y = q_x$, we incorporate the integrability constraints through a penalty term $(p_y - q_x)^2$ to the functional. This approach seems to work better than other approaches that try to strictly enforce the integrability constraints.[13] The parameter λ permits control of the trade-off between a "smoother" vector field that will enable better recovery of ϕ and a vector field that reaches lower visibility. The squaring of V and the particular form of the penalty term were chosen to produce simple Euler equations.

Adding a penalty term that represents the integrability constraint and squaring V results in the following functional:

$$I[p, q] = \int\int_\Omega (V^2(p, q) + \lambda(p_y - q_x)^2) \mathrm{d}x\mathrm{d}y. \qquad (17)$$

Equation (17) is in I_3 form, and its Euler equations are

$$-VV_p + \lambda(p_{yy} - q_{xy}) = 0,$$

$$-VV_q + \lambda(q_{xx} - p_{yx}) = 0. \qquad (18)$$

By discretizing Eqs. (18) we obtain the following iterative scheme:

$$p_{i,j}^{k+1} = \bar{p}_{i,j}^k - \frac{1}{2}\tilde{q}_{i,j}^k - \frac{1}{2\lambda}V(p_{i,j}^k, q_{i,j}^k)V_p(p_{i,j}^k, q_{i,j}^k),$$

$$q_{i,j}^{k+1} = \bar{q}_{i,j}^k - \frac{1}{2}\tilde{q}_{i,j}^k - \frac{1}{2\lambda}V(p_{i,j}^k, q_{i,j}^k)V_q(p_{i,j}^k, q_{i,j}^k),$$

where

$$\bar{p}_{i,j} = \frac{p_{i,j+1} + p_{i,j-1}}{2},$$

$$\bar{q}_{i,j} = \frac{q_{i+1,j} + q_{i-1,jx}}{2},$$

$$\tilde{p}_{i,j} = \frac{p_{i+1,j+1} + p_{i-1,j-1} - p_{i+1,j-1} - p_{i-1,j+1}}{4},$$

$$\bar{q}_{i,j} = \frac{q_{i+1,j+1} + q_{i-1,j-1} - q_{i+1,j-1} - q_{i-1,j+1}}{4}.$$

As initial conditions, we took an arbitrary vector field. The boundary conditions are described in Section 9. Note that the natural boundary conditions (16) in this case reduce to the integrability condition $p_y = q_x$ on the boundary. The above iterative scheme is simple, and although in our experiments it converged quickly, we expect more complicated methods such as multigrid methods to be more efficient.

7. RECOVERING HEIGHT FROM GRADIENT

A. Height-from-Gradient Problem
The height-from-gradient problem deals with the following problem: Given a vector field $\mathbf{F}(x, y)$, find a function $\phi(x, y)$ such that $\nabla \phi(x, y) = \mathbf{F}(x, y)$.

Note that the solution ϕ is not unique, since adding a constant term to ϕ will result in another solution to the problem. This problem can therefore be classified as an initial-value problem: Given an initial value at some location $\phi(x_0, y_0)$ and $\mathbf{F}(x, y)$, find $\phi(x, y)$ for all the region.

A simple solution to this problem is

$$\phi(x, y) = \phi(x_0, y_0) + \int_C \nabla \phi \cdot \mathbf{d}l, \qquad (19)$$

where C is a curve from (x_0, y_0) to (x, y). This method allows us to compute ψ completely once an initial value $\psi(x_0, y_0)$ is determined. The problem with Eq. (19) is that it is numerically unstable. A height value at some point would in the presence of noise depend on the integration path that was taken. It is better to find a best-fit surface $\phi^\star$ to ϕ. This can be accomplished by a variational calculus setting.[13] The variational approach to height from gradient is discussed in Subsection 7.B.

B. Variational-Calculus Setting for Height from Gradient
Given the vector field $\mathbf{F}(x, y) = (p(x, y), q(x, y))$ and a possible approximate solution $\phi^\star$, we wish to minimize the following functional:

$$\int\int_\Omega (\phi_x^\star - p)^2 + (\phi_y^\star - q)^2 \mathrm{d}x\mathrm{d}y. \qquad (20)$$

Calculating the Euler equation for functional (20) yields $\Delta \phi^\star = p_x + q_y$, where $\Delta \phi^\star$ is the Laplacian of $\phi^\star$: $\partial^2 \phi^\star/\partial x^2 + \partial^2 \phi^\star/\partial y^2$. This equation is a second-order elliptic partial differential equation called a Poisson equation. The Poisson equation has been widely studied, and

many procedures for numerical solutions exist. In our experiments we used two methods to solve this equation. One is a multigrid method, and the other is based on sine transforms and tridiagonal solutions.[14]

Once again, note that this equation does not uniquely specify a solution without further constraints. In fact, we can add any function h that satisfies $\Delta h = 0$ to the solution. For this particular problem, the natural boundary conditions are $(\phi_x^\star, \phi_y^\star) \cdot \mathbf{n} = (p, q) \cdot \mathbf{n}$, where $\mathbf{n}$ is the normal to the curve that represents the boundary. With these boundary conditions the solution is still not unique, since an arbitrary constant can be added to $\phi^\star$ without changing the functional. To get a unique solution, one can fix arbitrary height at some point.

8. EXPERIMENTAL RESULTS

When solving Eqs. (18) we have to specify a boundary condition and an initial value. The initial value is $p_{i,j}^0, q_{i,j}^0$ for all i, j in the domain. The boundary condition is the update rule from one iteration to the next along the boundary of the domain. Since we have no *a priori* knowledge of the boundary values, we will consider two methods for updating the boundary values between iterations.

If we have additional knowledge on the desired p and q, we may be able to incorporate this knowledge into the boundary condition. For example, if the desired pattern can be rolled up along the x and y axes to form the surface of a three-dimensional torus or donut, periodic boundary conditions can be used.

In periodic boundary conditions, the boundary value in the next iteration is taken from the computed values along the opposite boundary. For the case of an image I of size $N \times N$, the update rule is

$$p_{i,1}^{k+1} = p_{i,N-1}^k, \qquad p_{i,N}^{k+1} = p_{i,2}^k,$$
$$p_{1,i}^{k+1} = p_{N-1,i}^k, \qquad p_{N,i}^{k+1} = p_{2,i}^k,$$

and the boundary of $q_{i,j}^k$ is updated in a similar manner. Periodic boundary conditions will perform well for periodic shapes, but this is hardly the general case.

In the general case, if we knew p_y, q_x on the boundary, we could update the boundary values by integration:

$$p(1, j) = p(1, j - 1) + \int_{j-1}^{j} p_y(1, t)\mathrm{d}t,$$

$$p(N, j) = p(N, j - 1) + \int_{j-1}^{j} p_y(N, j)\mathrm{d}k,$$

$$p(i, 1) = p(i, 2) - \int_{2}^{1} p_y(i, t)\mathrm{d}t,$$

$$p(i, N) = p(i, N - 1) + \int_{N-1}^{N} p_y(i, t)\mathrm{d}t. \qquad (21)$$

The values of q can be similarly updated.

The integration can be numerically approximated by the trapezoidal rule (see Ref. 15, p. 190) as follows:

$$p(1, j) = p(1, j - 1) - \tfrac{1}{2}(p_y(1, j - 1) + p_y(1, j)),$$

$$p(N, j) = p(N, j - 1) + \tfrac{1}{2}(p_y(N, j - 1) + p_y(N, j)),$$

$$p(i, 1) = p(i, 2) - \tfrac{1}{2}(p_y(i, 2) + p_y(i, 1)),$$

$$p(i, N) = p(i, N - 1) + \tfrac{1}{2}(p_y(i, N) + p_y(i, N - 1)). \qquad (22)$$

We now turn to the problem of approximating the derivatives in Eq. (22). We can approximate $q_x(i, 1)$, $q_x(i, N)$, $i = 2 \ldots N - 1$ by the central-difference formula (see Ref. 15, p. 675). The values of $q_x(1, j)$, $q_x(N, j)$ can be approximated by forward or backward formulas. We can now use the fact that p and q satisfy $p_y = q_x$ on the boundary to compute p_y needed for the above computation.

The values at the corners $p(1, 1)$, $p(1, N)$, $p(N, 1)p(N, N)$, $q(1, 1)$, $q(1, N)$, $q(N, 1)$, $q(N, N)$ are subject to large numerical error, and we average them with values from neighboring pixels.

To evaluate the algorithm it would be desirable to synthesize moiré patterns whose optimal ϕ and ψ are known. We could then compare the optimal ϕ and ψ with the functions found by the minimization process.

Finding optimal ϕ and ψ for arbitrary moiré patterns requires exhaustive search over a function space. Such search is in the general case clearly impractical. However, for certain moiré patterns, optimal synthesis can be computed without exhaustive search. An example of such moiré patterns is linear patterns.

From the structure of the optimality criterion it is clear that the optimal $\nabla \phi$ should be constant throughout the image. The optimal ϕ should therefore be a linear image.

To find the optimal ϕ we proceed as follows. Every linear function ϕ is characterized by two parameters $(p, q) = \nabla \phi$. Finding the optimal ϕ reduces in this case to evaluating the performance criteria over $\mathbb{R}^2$. The optimal ϕ is not unique, since the performance criteria is symmetric with respect to reflection around the lines $x = 0$, $y = 0$, $y = x$, $y = -x$. In other words, $V(p_0, q_0) = V(p_0, -q_0) = V(-p_0, q_0)$, and so on.

We then check the solution found by the iterative scheme. Starting from an initial condition of $(p, q) = (0, 1)$ and using the boundary conditions (22), we arrive at the solution whose gradient is identical to one of the optimal gradient. The initial values and the values at iteration 20 are shown in Fig. 5. The dotted vectors represent the gradient of the original linear image, and the solid vectors represent the computed $\nabla \phi$.

It is interesting to start with two functions, create a superposition, and feed this superposition to the iterative procedure. In general, the solution will not be the same as the two original functions. The reason for this is that the superposition we started with was probably not optimal or that the algorithm converges to a different local minimum.

However, we succeeded in calculating the two original functions in the following case. We start with two ellipses whose centers are shifted along the x axis. The equations for two such ellipses are

$$\frac{(x-s)^2}{a^2} + \frac{y^2}{b^2} = h^2,$$

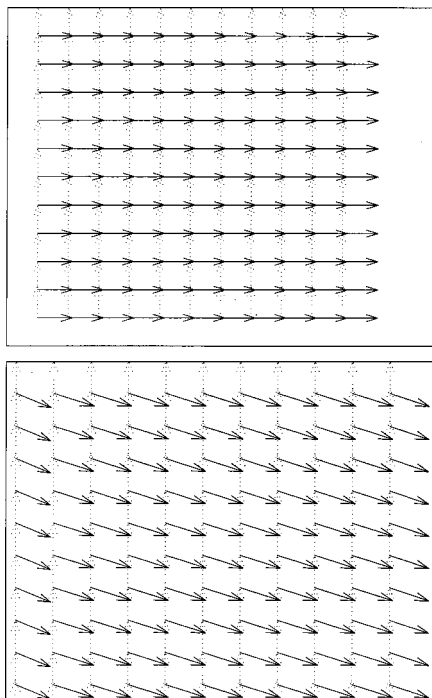$$\frac{(x+s)^2}{a^2} + \frac{y^2}{b^2} = k^2.$$

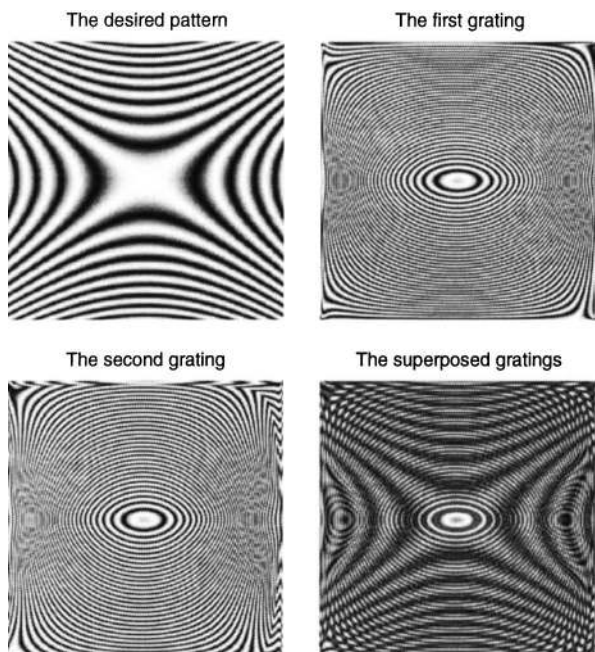Fig. 5. Linear moiré, initial condition (top), and values after 20 iterations (bottom).

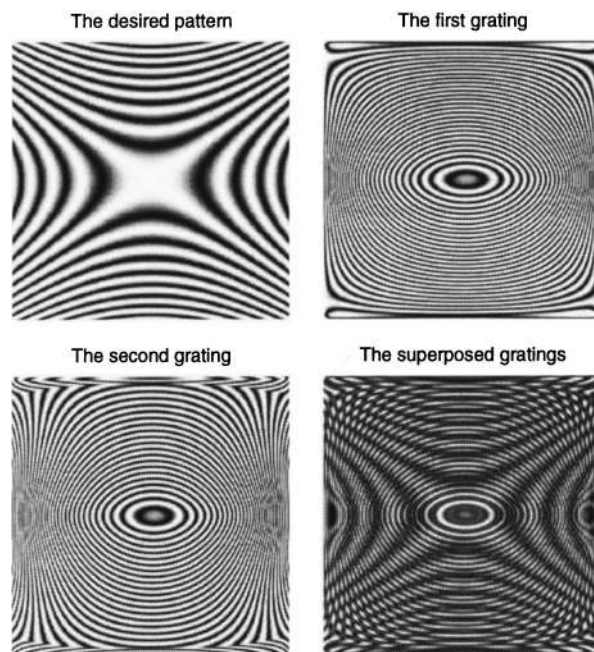Fig. 6. Hyperbolic patterns: results for natural boundary conditions.

Fig. 7. Hyperbolic patterns: results for periodic boundary conditions
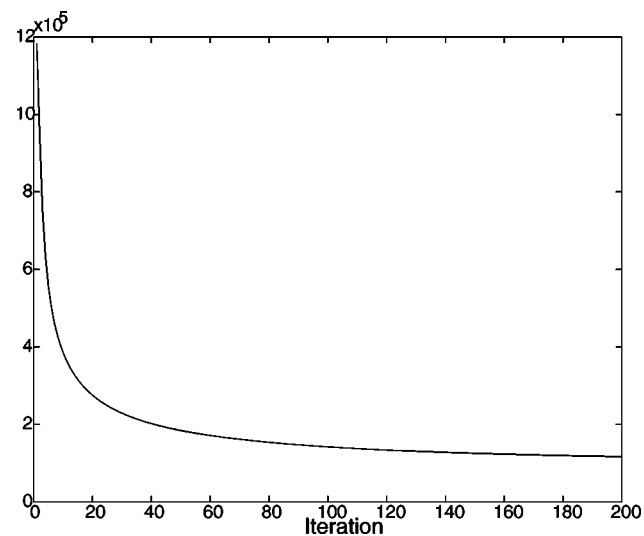
Fig. 8. Decrease in functional value.

The indicial equation is $h - k = p$. By elimination of h and k from these equations and after some rearrangements, the following equation is obtained,[5]

$$\frac{4x^2}{a^2 p^2} - \frac{y^2}{(b^2 p^2 / 4) - b^2 s^2} = 1, \tag{23}$$

which represents a hyperbola parameterized by p.

Indeed, when we used the synthesis algorithm to produce hyperbolic moiré patterns, such ellipses were found. In Fig. 6 results are shown for natural boundary conditions, and in Fig. 7 results are shown for periodic boundary conditions. The iterative process converged fast in our experiments. Usually after approximately 200 iterations there was no apparent change in the images. The minimized functional values for the above experiment are plotted in Fig. 8.

The parameter λ in Eq. (17) allows control of the "smoothness" of the solution. The results in Fig. 9 were obtained for $\lambda = 200$. Compare these images with Fig. 10, which was obtained for $\lambda = 1000$.

A face image $g(x, y)$ and its raised cosine periodic profile $0.5 \cos[2\pi f g(x, y)] + 0.5$ are shown in Figs. 11 and 12. The computed ϕ and ψ and the superposition image are shown in Figs. 13 and 14. The periodic profile of the face image in Fig. 15 is shown in Fig. 16. The computed ϕ and ψ are shown in Fig. 17, and the superposition image is shown in Fig. 18.

In our scheme for moiré synthesis, the superposition image consists of low-frequency and high-frequency com-

Fig. 11.   Face image 1.
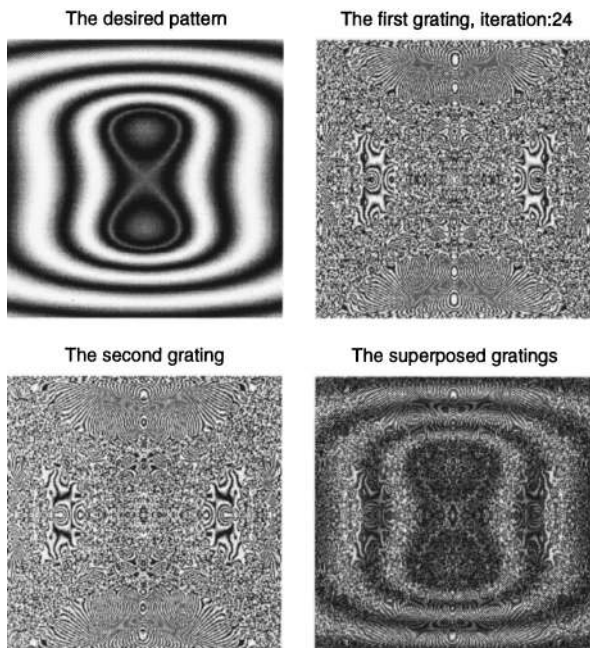

Fig. 12.   Periodic profile of Fig. 11.

Fig. 9.   Result for $\lambda = 200$.
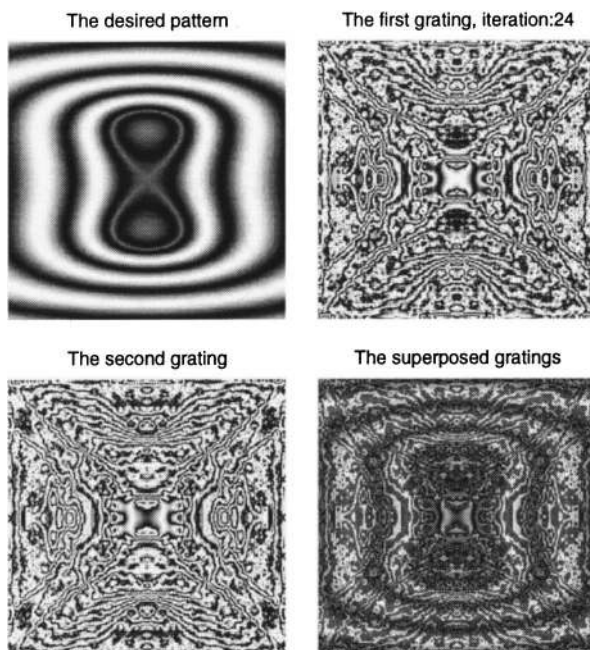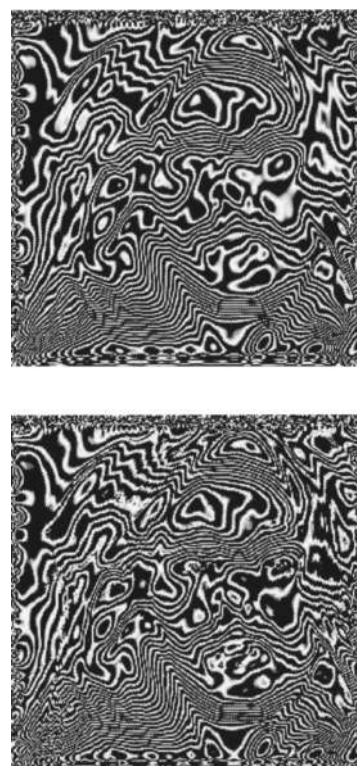

Fig. 10.   Result for $\lambda = 1000$.

Fig. 13. ψ and ϕ computed for Fig. 12.

ponents. The low-frequency components represent the desired pattern and zero-order moiré patterns (if they exist). The high-frequency components represent the (1, 1) moiré terms that were "pushed" outside the visibility circle and higher-order moirés. We can therefore apply low-pass filtering to the superposition image to enhance the desired pattern. An example of applying simple low-pass filtering is shown in Fig. 19. Compare the reconstructed image (Fig. 19) with the original pattern (Fig. 16).

If, in addition, the periodic profile does not have a dc term, zero-order moiré patterns do not exist, and hence we expect better reconstruction.
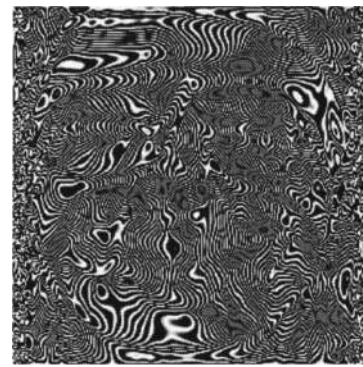

Fig. 14.   Superposition of the images in Fig. 13.
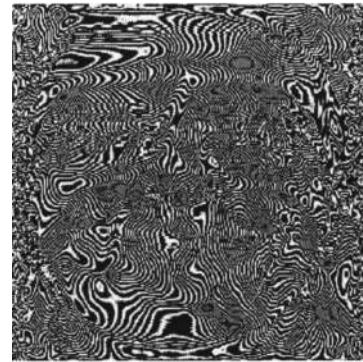

Fig. 15.   Face image 2.

Fig. 16.   Periodic profile of Fig. 15.

a

b

Fig. 17. ψ and ϕ computed for Fig. 16.

Fig. 18.   Superposition of the images in Fig. 17.

Fig. 19.   Low-pass filtering applied to Fig. 18.

9. CONCLUSIONS

The suggested visibility criteria seem to produce good results especially in simple cases such as Fig. 6. In more-complicated images (such as the face images), the optimization algorithm seems to converge to a local minimum, and the final result depends on the initial conditions. Although in the general case the boundary condition is unknown, experimental results show that this affects only solution pixels near the boundary.

The results of this work suggest another application that might have been overlooked. If the desired pattern is smooth, the two original images bear little or no resemblance to the desired pattern. The desired pattern is created by the nonlinear superposition of the two images.

Moiré pattern synthesis may then be used for some sort of visual cryptography. Instead of transmitting the image on an unsecured channel, it is possible to transmit two images that create a moiré pattern of the desired pattern. However, note that for nonsmooth images such as the face image in Fig. 15, areas of discontinuities in ψ and ϕ may indicate the location of the boundary of the face.

A method for visual cryptography for binary images has been proposed in Ref. 16 that allows perfect reconstruction, but the reconstructed image is half the resolution of the transmitted images. Extending this method to gray-level images will require transmitting two very large images. In moiré synthesis, perfect reconstruction is not possible. However, as is seen in the previous examples, it is often easy to recognize the pattern from the superposition. Our performance criteria were designed for visibility of the moiré patterns. Other applications, such as visual cryptography, certainly require different performance criteria. The optimization scheme, however, may remain the same.

ACKNOWLEDGMENTS

Address correspondence to Guy Lebanon at the address on the title page or by e-mail, lebanon@cs.cmu.edu.

REFERENCES

1. H. Giger, "Moirés," Comput. Math. Appl. **12**, 329–361 (1986).
2. Lord Rayleigh, "On the manufacture and theory of diffraction-gratings," Philos. Mag. **81**, 81–93 (1874). Published also in G. Indebetouw and R. Czarnek, *Selected Papers on Optical Moiré and Applications* Vol. 64 of SPIE Milestone Series (SPIE Press, Bellingham, Wash., 1992), pp. 3–15.
3. O. Bryngdahl, "Characteristics of superposed patterns in optics," J. Opt. Soc. Am. **66**, 87–94 (1976).
4. I. Amidror, *The Theory of the Moiré Phenomenon* (Kluwer Academic, Dordecht, The Netherlands, 2000).
5. M. Wasserman, G. Oster, and C. Zwerling, "Theoretical interpretations of moiré patterns," J. Opt. Soc. Am. **54**, 169–175 (1964).
6. R. Courant and D. Hilbert, *Methods in Mathematical Physics* (Interscience, New York, 1953), Vol. 1.
7. W. K. Pratt, *Digital Image Processing*, 2nd ed. (Wiley, New York, 1991).
8. R. Ulichney, *Digital Halftoning* (MIT Press, Cambridge, Mass., 1987).
9. M. Levine, *Vision in Man and Machine* (McGraw-Hill, New York, 1985).
10. M. Taylor, "Visual discrimination and orientation," J. Appl. Opt. **53,** 763–765 (1963).
11. G. Thomas and R. Finney, *Calculus and Analytic Geometry*, 9th ed. (Addison-Wesley, Reading, Mass., 1996).
12. H. Sagan, *Introduction to the Calculus of Variations* (Dover, New York, 1969).
13. B. K. P. Horn and M. J. Brooks, "The variational approach to shape from shading," Comput. Vis. Graph. Image Process. **2**, 174–203 (1986).
14. G. Strang, *Introduction to Applied Mathematics* (Wellesley–Cambridge Press, Wellesley, Mass., 1986).
15. R. L. Burden and J. D. Faires, *Numerical Analysis*, 6th ed. (Brooks Cole, Pacific Grove, Calif., 1997).
16. M. Naor and A. Shamir, "Visual cryptography," in *EUROCRYPT 1994, Lecture Notes in Computer Science* (Springer-Verlag, New York, 1995), Vol. 950, pp. 1–12.