

Single-Band Amplitude Demodulation of Müller-Lyer Illusion Images

Vicente Sierra-Vázquez¹ and Ignacio Serrano-Pedraza²

¹Universidad Complutense, Madrid

²University of Newcastle

The perception of the Müller-Lyer illusion has previously been explained as a result of visual low band-pass spatial filtering, although, in fact, the illusion persists in band-pass and high-pass filtered images without visible low-spatial frequencies. A new theoretical framework suggests that our perceptual experience about the global spatial structure of an image corresponds to the amplitude modulation (AM) component (or its magnitude, also called envelope) of its AM-FM (alternatively, AM-PM) decomposition. Because demodulation is an ill-posed problem with a non-unique solution, two different AM-FM demodulation algorithms were applied here to estimate the envelope of images of Müller-Lyer illusion: the global and exact Daugman and Downing (1995) AMPM algorithm and the local and quasi-invertible Maragos and Bovik (1995) DESA. The images used in our analysis include the classic configuration of illusion in a variety of spatial and spatial frequency content conditions. In all cases, including those of images for which visual low-pass spatial filtering would be ineffective, the envelope estimated by single-band amplitude demodulation has physical distortions in the direction of perceived illusion. It is not plausible that either algorithm could be implemented by the human visual system. It is shown that the proposed second order visual model of pre-attentive segregation of textures (or "back-pocket" model) could recover the image envelope and, thus, explain the perception of this illusion even in Müller-Lyer images lacking low spatial frequencies.

Keywords: visual optical illusions, amplitude demodulation, AMPM algorithm, DESA, Müller-Lyer

La percepción de la ilusión de Müller-Lyer ha sido explicada como resultado del filtrado visual paso-bajo de las imágenes en las que aparece, aunque, de hecho, la ilusión se percibe en imágenes paso-banda y paso-alto carentes de bajas frecuencias espaciales. Una nueva manera de pensar acerca del procesamiento visual espacial sugiere que la percepción de la estructura espacial global de una imagen se corresponde con el componente de amplitud modulada (AM) o envolvente resultante de su descomposición AM-FM (o, alternativamente, de su descomposición AM-PM). En este trabajo, la envolvente de imágenes de la ilusión de Müller-Lyer se estimó mediante dos algoritmos de demodulación: el algoritmo AMPM de Daugman y Downing (1995) y DESA de Maragos & Bovik (1995). Las imágenes de Müller-Lyer utilizadas presentan la configuración clásica de la ilusión en diferentes versiones espaciales y con diferente contenido en frecuencia espacial. Para cada una de las imágenes utilizadas, incluidas aquellas en las que su filtrado paso-bajo es inútil para obtener su estructura global, la envolvente estimada mediante la demodulación de la amplitud presenta distorsiones físicas que se corresponden con la ilusión percibida. Es poco plausible que el sistema visual humano implemente cualquiera de los dos algoritmos utilizados. Sin embargo, se muestra que el modelo de mecanismos visuales de segundo orden propuesto para la segregación preatencional de la textura puede recuperar la envolvente de los estímulos visuales, explicándose así la percepción de la ilusión de Müller-Lyer aún en imágenes carentes de bajas frecuencias espaciales.

Palabras clave: ilusiones visuales, demodulación de la amplitud, algoritmo AMPM, DESA, Müller-Lyer

This research was supported by grant #BSO2002-00192 from Ministerio de Ciencia y Tecnología (Spain). M.A. García-Pérez wrote the FORTRAN code to generate the balanced-square image (Figure 1B) and postscript routines to plot 3D graphics. We are also grateful to two anonymous reviewers for valuable comments and helpful suggestions. Preliminary results were presented at the 34th European Mathematical Psychology Group Meeting held in Madrid, August 26-30, 2003 and reported in abstract form by Sierra-Vázquez & Serrano-Pedraza (2003).

Address correspondence to: Dr. V. Sierra-Vázquez, Departamento de Psicología Básica I, Facultad de Psicología, Universidad Complutense de Madrid, Campus de Somosaguas, 28223 Madrid, Spain. Phone: +34 913 943 144. Fax: +34 913 943 189. E-mail: vicente@psi.ucm.es

The Müller-Lyer illusion or “confluxion paradox” is perhaps the most famous of all geometrical visual optical illusions (Boring, 1942). It is an illusion of extent where two line segments of physically equal length are perceived as unequal: a straight line (or shaft) between inward-pointing arrowheads looks longer than an identical shaft between outward-pointing arrowheads (Figure 1A). The effect not only occurs in the complete or incomplete original figures of illusion (Müller-Lyer, 1889, 1896; Brentano, 1892), but also in a variety of configurations in which the shafts are bounded by contours other than arrowheads or arrow-feathers (Heymans, 1896; Robinson, 1972, his Figures 2.2 to 2.6 and 2.8 to 2.10). The phenomenon is so surprising that a large variety of explanations of this illusion have been given since the first one offered by Müller-Lyer himself, which was based on “contrast” and “confluxion” principles, and other classic explanations collected by Boring and Robinson.

Among the modern influential explanations in the field of early visual processing, Ginsburg (1978, 1986) suggested that the Müller-Lyer illusion (as well as other veridical and non-veridical visual phenomena) is physically carried by the low-spatial-frequency range of the Fourier spectrum of the image (i.e., in the coarse spatial scales) and, thus, its perception can be explained as the result of a visual low band-pass frequency filtering. In spite of the difficulties enumerated by Gregory (1990), Ginsburg’s theory has some advantages over other explanations: It does not assume *ad hoc* processes for this particular illusion, the proposed linear filtering mechanisms are consistent with the known functional architecture of the human visual system, and it is formulated in computable terms. In addition, results reported by Carrasco, Figueroa, and Willen (1986) offered experimental support for this theory: According to Ginsburg’s hypothesis, the illusion in one of the Brentano forms (Brentano, 1892, his Figure 9) decreases after adaptation to a vertical grating of low spatial frequency, but seems unaffected by adaptation to horizontal or high-frequency vertical gratings.

However, several studies have shown that the illusion persists in a number of images containing few or no low spatial frequencies and, because of that, these images leave the first-order psychophysical channels tuned to low spatial frequencies theoretically inactive. To illustrate this, Figure 1B shows an image made from balanced squares (García-Pérez, 1991), similar to the illusion image of Carlson, Moeller, and Anderson (1984) composed of balanced dots, and Figure 1C shows a high-pass filtered image: In both cases, the upper shaft appears longer than the lower one

although the images lack energy in the low-spatial frequency range (see the corresponding relative energy distribution). In addition, it can be demonstrated that the information about the illusion may be in spectral (spatial) bands other than low spatial frequency ranges. Figure 1D depicts a second-order version of illusion where a low-pass filtered Müller-Lyer image is the envelope of a vertical sinusoidal grating of high spatial frequency. Because of the modulation theorem (Bracewell, 1978), the band-limited spectrum of the envelope is shifted along the u axis of the Fourier plane (and so is its energy distribution). Digital alterations of spectra seem to demonstrate, according to Skottun (2000), that the illusion is not the result of low-spatial filtering because it is perceived in hybrid images¹ and phase-only images generated from flat amplitude spectra (Figure 1E). Finally, Gregory (1990) claims that, theoretically, Ginsburg’s explanation (as a type of physiological lateral inhibition theory) could not explain the appearance of the illusion in the Brentano figure without shafts (Brentano, 1892, his Figures 5 and 6; Gregory, 1990, his Figure 9.7; see also our Figure 1F). Therefore, the perceptual appearance of the Müller-Lyer illusion needs a different kind of explanation.

In the field of early spatial vision, a new theoretical framework suggests that:

1. Our perceptual experience about the global spatial structure of an image corresponds to the amplitude modulation (AM) component of the AM-FM decomposition of the given image (Daugman & Downing, 1995); and
2. The amplitude envelope may be perceptually retrieved (demodulated) by some nonlinearity of the visual system (Schofield & Georgeson, 2003).

In this work, the former suggestion is investigated by estimating the envelope of variants of the classic Müller-Lyer configuration with two different AM-FM demodulation algorithms; in addition, it will be shown that amplitude demodulation could be carried out by second-order visual mechanisms similar to those that have been proposed to explain pre-attentive texture segregation, providing a psychophysical basis to explain also the perception of this illusion.

Mathematical Background

According to a single-band modulation model, any narrowband image can be expressed *in the spatial domain* as the combination of a smoothly varying two-dimensional spatial

¹ A hybrid image (Tadmor & Tolhurst, 1993) or image with a hybrid spectrum (Skottun, 2000) is created by computing the inverse Fourier transform of a spectrum made by taking the Fourier amplitude spectrum from one image and the Fourier phase spectrum from another. (See also Oppenheim & Lim, 1981).

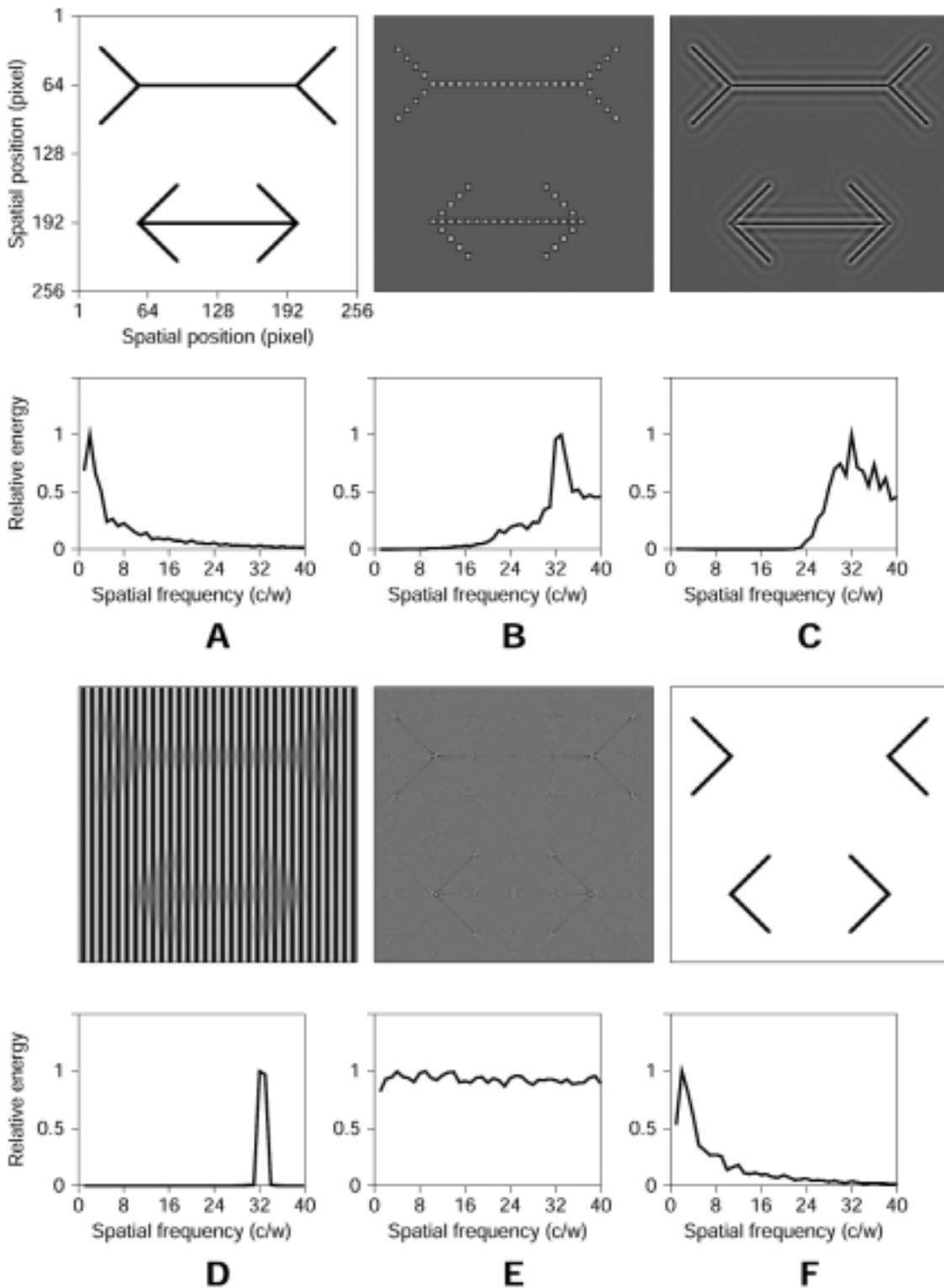


Figure 1. Standard image and variants eliciting the Müller-Lyer illusion. A: Conventional standard version of the Müller-Lyer configuration. B: Balanced-square image (García-Pérez, 1991). C: High-pass-filtered version of the conventional image. The sharp edges of the isotropic ideal high-pass filter were smoothed by a cycle of a raised cosine (radial spatial frequency cut-off at 21 c/w). D: A second-order version of the Müller-Lyer illusion. The image is a vertical sinusoidal grating (spatial frequency equal to 32 c/w) as carrier whose contrast is modulated by a low-pass filtered version of the standard configuration of Müller-Lyer image. E: Phase-only image generated from a flat amplitude spectrum (after Skottun, 2000). F: Müller-Lyer arrow-heads, drawn without the shafts that connect them (Brentano, 1892; Gregory, 1990). The normalized energy spectrum averaged across all orientations appears under each image.

envelope (the AM component) and a fine spatial structure (frequency modulated 2D carrier or FM 2D wave). In the combination, the squared local amplitude represents local energy, and the local phase represents local spatial structure. This single-band model is not appropriate for real-world images because they are broadband and spatially non-stationary signals, but its generalization is straightforward considering broadband signals as a sum of narrow-band signals obtained via spectrally disjoint band-pass filtering (Havlicek & Bovik, 2000). Both the concept of envelope (or local amplitude) and the concept of spatial frequency modulation (or, alternatively, local phase) can only be defined with respect to some complex extension of the given real image.

Mathematically, a real 2D signal s , $-1 \leq s(x,y) \leq 1$, can be completely represented, in the spatial domain, as the real part of the complex modulation product of a spatially non-stationary amplitude and a single complex 2D FM carrier, that is,

$$s(x,y) = \text{Re}\{a(x,y)c(x,y)\}, \quad (1)$$

where the local amplitude a is the AM component or AM function (its magnitude, $|a| \geq 0$, is called the image envelope) and c , $c(x,y) = \exp\{j[2\pi(u_c x + v_c y) + \varphi(x,y)]\}$, ($j = \sqrt{-1}$), is a complex 2D FM modulated carrier, φ being the local phase, the PM component or PM function, and u_c and v_c , respectively, being the carrier center spatial frequencies. The local amplitude and local phase are, respectively, measures of quantitative and qualitative information of the image: Relative maxima of local amplitude indicate the location of spatial features and specific values of local phase indicate the class of features (Morrone & Burr, 1988; Sierra-Vázquez & García-Pérez, 1995). In addition, given the smooth spatial variation of the local amplitude, the AM component conveys information about the global spatial structure of image.

By definition, the gradient vector (Ω_2, Ω_1) , whose components are

$$\begin{aligned} \Omega_1(x,y) &= 2\pi v_c + \frac{\partial \varphi(x,y)}{\partial y}, \\ \Omega_2(x,y) &= 2\pi u_c + \frac{\partial \varphi(x,y)}{\partial x}, \end{aligned} \quad (2)$$

gives the spatial frequency modulation function or FM function, being $\Omega_1(x, y)$, $\Omega_2(x, y)$, in turn, the vertical and horizontal instantaneous angular spatial frequencies (Maragos, & Bovik, 1995), in radians per unit of length, in the position (x,y) . Partial derivatives of phase function, Ω_1 , Ω_2 , are also referred to as *emergent image frequencies* (Bovik, Gopal, Emmoth, & Restrepo [Palacios], 1992). Putting together local amplitude and local phase in a complex-valued function Z , Equation (1) can be rewritten as

$$s(x,y) = \text{Re}\{Z(x,y)C(x,y)\}, \quad (3)$$

with $C(x,y) = \exp\{j[2\pi(u_c x + v_c y)]\}$ being the periodic carrier and $Z(x,y) = A(x,y)\exp[j\varphi(x,y)]$ being the complex amplitude,

where A is the AM component and φ is the PM component. Hence, the given 2D signal can be alternatively represented by the real part of the modulation product of a single, periodic complex carrier C and a complex phasor Z that specifies, over the spatial domain, its amplitude and its phase modulations (Daugman & Downing, 1995). Hereafter, the terms AM-FM and AM-PM are used interchangeably as FM and PM functions are directly related by a linear operation (differentiation or integration) and so, either one can be obtained from the other.

Demodulation of a given image is the process of estimating its AM and FM components (alternatively, its carrier frequencies and their AM and PM components). Specifically, single-band amplitude demodulation of a given image is the process of calculating the AM component of a band-pass version of the image. However, demodulation is an ill-posed problem because, for any real signal, there is an unlimited number of combinations of AM and FM components, the real part of whose modulation product is just the given signal (for 1D signals, see Vakman, 1972, and Loughlin & Tacer, 1996; for 2D signals, see Havlicek & Bovik, 2000). The aim of image demodulation algorithms is to compute effectively the inherent carrier, amplitude, and frequency (or phase) modulations of a given image. To limit the field of possible solutions, it is common practice to impose requirements such as small estimation errors for known components, complete reconstruction of the original image from its estimated components, or physical meaningfulness of the estimated components.

It is not guaranteed that different algorithms will give identical or similar estimates of modulation components when applied to the same image. Because of that, here, we apply two procedures to estimate the AM component of images of the Müller-Lyer illusion and then we compare the results obtained with each one. Each of the algorithms used here belongs to one of the two general approaches defined by Havlicek and Bovik (2000) to tackle the image demodulation problem. One of them is Maragos and Bovik's (1995) discrete energy separation algorithm (DESA) based on a multidimensional extension of the discrete 1D Teager-Kaiser non-linear energy-tracking operator (TKEO; Kaiser, 1990; Maragos, Kaiser, & Quatieri, 1993). Given the 2D signal s , the goal of DESA is to estimate $|a|$, Ω_1 , and Ω_2 such that,

$$\begin{aligned} \hat{s}(x,y) &= |a(x,y)| \cos\left[\int_0^x \Omega_2(p,q) dp + \int_0^y \Omega_1(p,q) dq\right], \\ |a(x,y)| &\geq 0, \end{aligned} \quad (4)$$

where \hat{s} stands for the signal estimated from modulation components. The other procedure used here is the Daugman and Downing's (1993, 1995) anisotropic AMPM algorithm based on a 2D extension of classic demodulation algorithms for 1D signals in acoustics (Hartmann, 1998) and which became popular because of its practical application for the

automatic identification of a person using a PM code of the texture of the iris image (Daugman & Downing, 1993). Given the signal s , the goal of Daugman and Downing's AMPM algorithm is to find the pair of carrier spatial frequencies (u_c, v_c) , and the complex phasor Z , with modulus A , and argument φ , such that

$$\begin{aligned}\hat{s}(x,y) &= 2A(x,y)\cos[2\pi(u_c x + v_c y) \varphi(x,y)], \\ A(x,y) &\geq 0, \\ -\pi &< \varphi(x,y) \leq \pi\end{aligned}\quad (5)$$

It should be noted that the cosine term on the right-hand side of Equation 5 is an FM carrier like its counterpart in Equation 4.

Let s be the contrast function corresponding to a digital image I , $0 \leq I(m,n) \leq 256$, $1 \leq m \leq M$, $1 \leq n \leq N$. Because \hat{s} may take negative values, reconstruction of the digital image from its AM-FM decomposition, \hat{I} , would be given by

$$\hat{I}(m,n) = I_0[1 + k\hat{s}(m,n)], \quad -1 \leq \hat{s}(m,n) \leq 1, \quad 0 \leq k \leq 1, \quad (6)$$

where $I_0=128$, k is the carrier contrast, and $\hat{s}(m,n)$ is a discrete version of the estimated signal $\hat{s}(x,y)$.

Methods

Images

The synthetic images (256×256 pixels) used here consisted of the standard Müller-Lyer configuration² (Figure 1A) and some variants of it: (a) an image made from balanced squares (Figure 1B) (García-Pérez, 1991); (b) a high-pass-filtered version of the original figure (Figure 1C); (c) a second-order version of the Müller-Lyer configuration, which has a vertical sinusoidal grating carrier whose contrast is modulated by a low-pass filtered conventional Müller-Lyer configuration (Figure 1D); (d) the phase-only image produced by combining the phase spectrum of the conventional figure with a flat amplitude spectrum (Skottun 2000, his Figure 5), similar to the Brentano-Gregory figure although less conspicuous (Figure 1E); and (e) the Brentano-Gregory version with fins only (Brentano, 1892, his Figures 5 and 6; Gregory, 1990, his Figure 9.7) (Figure 1F). The standard conventional configuration used here is often found in textbooks on perception (Goldstein, 1984; Sekuler & Blake, 1994); its variants have been chosen because of the appearance of illusion in them is a challenge for current explanations (as described in the Introduction). To eliminate

the ringing caused by ideal filtering and at the same time to circumvent the problem of residual low spatial frequencies addressed by García-Pérez (1991), a cycle of a raised cosine was used to smooth the sharp 1D profile of the modulation transfer functions (MTFs) of digital high-pass and low-pass filters used to create the high-pass and second-order versions of the illusion figure, respectively. The panel under each image in Figure 1 shows the energy spectrum averaged (and normalized) across all orientations. As shown, images B, C, and D are numerically devoid of energy in low spatial frequencies below 16, 21, and 30 cycles per image width (c/w) respectively.

Calibration of the Algorithms and Procedure

Custom software was written whose performance was calibrated by comparing the AM-FM estimates with known components of synthetic AM-FM signals. First, we replicated the results obtained by Maragos and Bovik (1995) and Daugman and Downing (1995) with their own test signals (an AM-FM oriented fringe used by Maragos & Bovik, 1995, their Figure 1 (a), and a number of narrow band-pass 2D textures in Daugman & Downing, 1995, their Figures 1 to 6). Second, a spatial FM carrier whose amplitude was modulated by a low-pass filtered letter C was used as further test image (see the Appendix). The implementation of DESA faces two main problems: Its input has to be a narrow-band signal and its output contains outlier estimates (including negative amplitude estimates) because of local numerical singularities; on the other hand, direct application of AMPM to broadband images does not work for some of them. To go around these difficulties, we pre-filtered the image with a band-pass filter before applying the algorithm, and then smoothed the raw result with a low-pass filter (see the flowchart in Figure 2A). Parameters of the band-pass pre-filter were chosen so that the selected band was appropriate to the spatial-frequency content of images and, thus, the filter had a center spatial frequency of 32 c/w and a relative bandwidth (full width at half-height) of one octave. Mathematical characterization of test image and technical details of the algorithms are given in the Appendix. Numerical results of this calibration indicated that highly accurate AM (and FM or PM, under some conditions) estimates can be obtained with both algorithms, and the image reconstructed from the components obtained by application of the AMPM algorithm is almost an exact copy of the original (Figure A2, column B, in the Appendix). In the application of the demodulation algorithm, all digital images were converted to a zero-mean floating-point 2D signal. Custom software was written in FORTRAN 77 using NAG (1991) subroutines for FFT; use of Matlab (MathWorks Ltd.) gave identical results.

² The original Müller-Lyer figure of 1889 has vertically oriented shafts, one part above the other (Müller-Lyer, 1889, his Figures 2a to 2g).

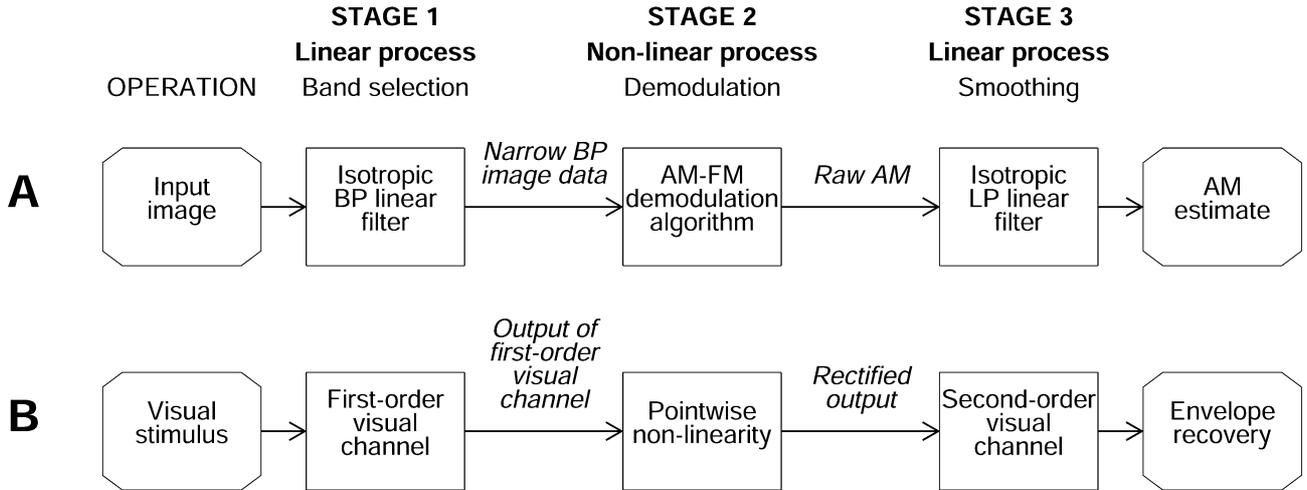


Figure 2. A: Flowchart of the single-band AM-FM (or AM-PM) demodulation used in Figures 3, 4, and 5. BP, band-pass. LP, low-pass. B: Flowchart of the psychophysical model used in Figure 6.

Quantification of the Illusion

The physical presence of the illusion was determined by measuring the distance between the two maxima in the respective estimated AM profiles at the actual location of the shaft in the input image.

Results

Figure 3 shows, for different versions of Müller-Lyer images (first column), the narrow band-pass filtered images (second column), amplitude envelopes estimated with DESA (third column), and amplitude profiles along the location of each shaft (fourth column). Estimated carriers, PM and AM estimates, and the AM profiles resulting from the application of the AMPM algorithm to the same band-pass images are shown in Figure 4 (input image and band-pass filtered images are not shown in this figure). Note that in Figures 3 and 4, the AM component physically carries the global structure of the image (i.e., the structure of the image as a whole) and also carries information about contrast, whereas the PM component (in Figure 4) picks up information about the fine spatial structure of the input image, despite the fact that its estimation depends on the choice of the values for the carrier spatial frequencies (see the Appendix). Both algorithms yield similar qualitative

results in the estimated AM component. In all cases, distances between maxima in the estimated AM component are longer for the upper shaft (limited by inward arrowheads) than in the lower (limited by outward arrowheads). These results suggest that the illusion is physically present in the AM component of single-band images of the standard and derived configurations, including those without low spatial frequencies (rows B, C, and F in Figures 3 and 4). Note the physical presence of the illusion also in the AM profiles of Brentano-Gregory and Skottun configurations (rows D and E, respectively, in Figures 3 and 4).

Chimaeric Images of Conventional Müller-Lyer Configuration

Our results show that the illusion is carried in the AM component, but closer inspection of the geometrical structures of PM component estimates suggests that, perhaps, the illusion is also in these components. If the illusion were exclusively conveyed by the AM component of the images, then it would persist when the AM component is combined with any carrier. To show that this is the case, we applied the AM-PM decomposition to generate chimaeric images³ of the conventional Müller-Lyer configuration. Like single-band auditory chimaeras (Smith, Delgutte, & Oxenham, 2002), a single-band visual chimera is the image resulting

³ Do not mistake chimaeric images for the hybrid images defined above: The former are made from the envelope and FM carrier in the spatial domain; the latter are made from the amplitude and phase spectra in the Fourier domain and taking the inverse Fourier transform.

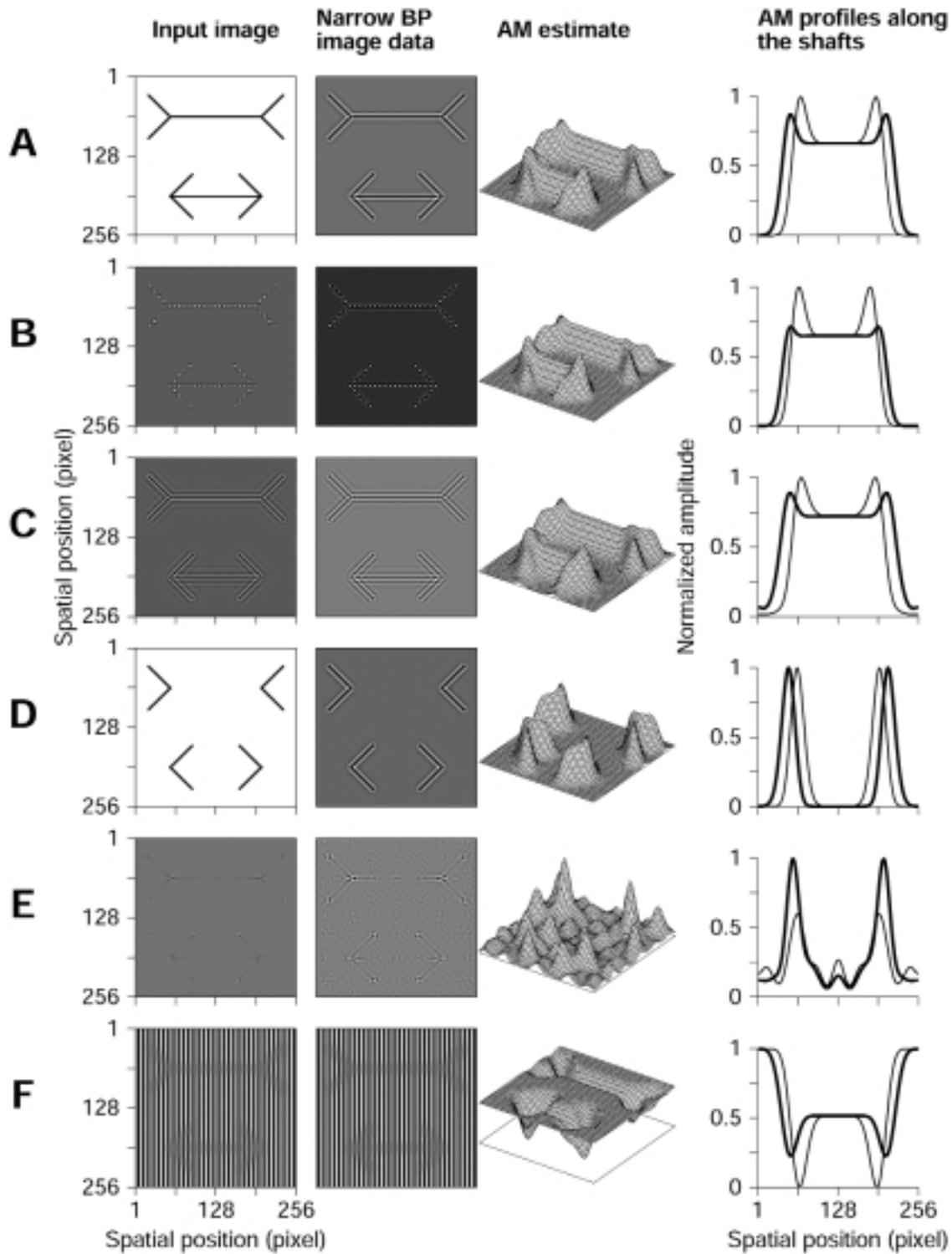


Figure 3. Amplitude envelopes of Müller-Lyer images estimated with Maragos and Bovik’s (1995) DESA and physical presence of the illusion. Input images (first column) are the images of Figure 1. A: Image of conventional figure. B: Balanced-square image. C: High-pass filtered version of the conventional figure. D: Brentano-Gregory image. E: Phase-only Skottun image. F: Image of a second-order version of the illusion. Spectral (spatial) narrow band-pass images (second column) were created with an isotropic band-pass filter whose MTF radial profile is a cycle of a raised cosine (radial center spatial frequency at 32 c/w and spatial frequency bandwidth of 1 octave (full-width at half height)). The radial profile of the MTF of the isotropic low-pass smoothing filter was a Gaussian function with a standard deviation of 8 c/w. AM estimates (third column) were normalized to unity. AM profiles (fourth column) along the shafts have their relative amplitudes unaltered within each panel (thick line, AM profile along the upper shaft; thin line, profile along the lower shaft).

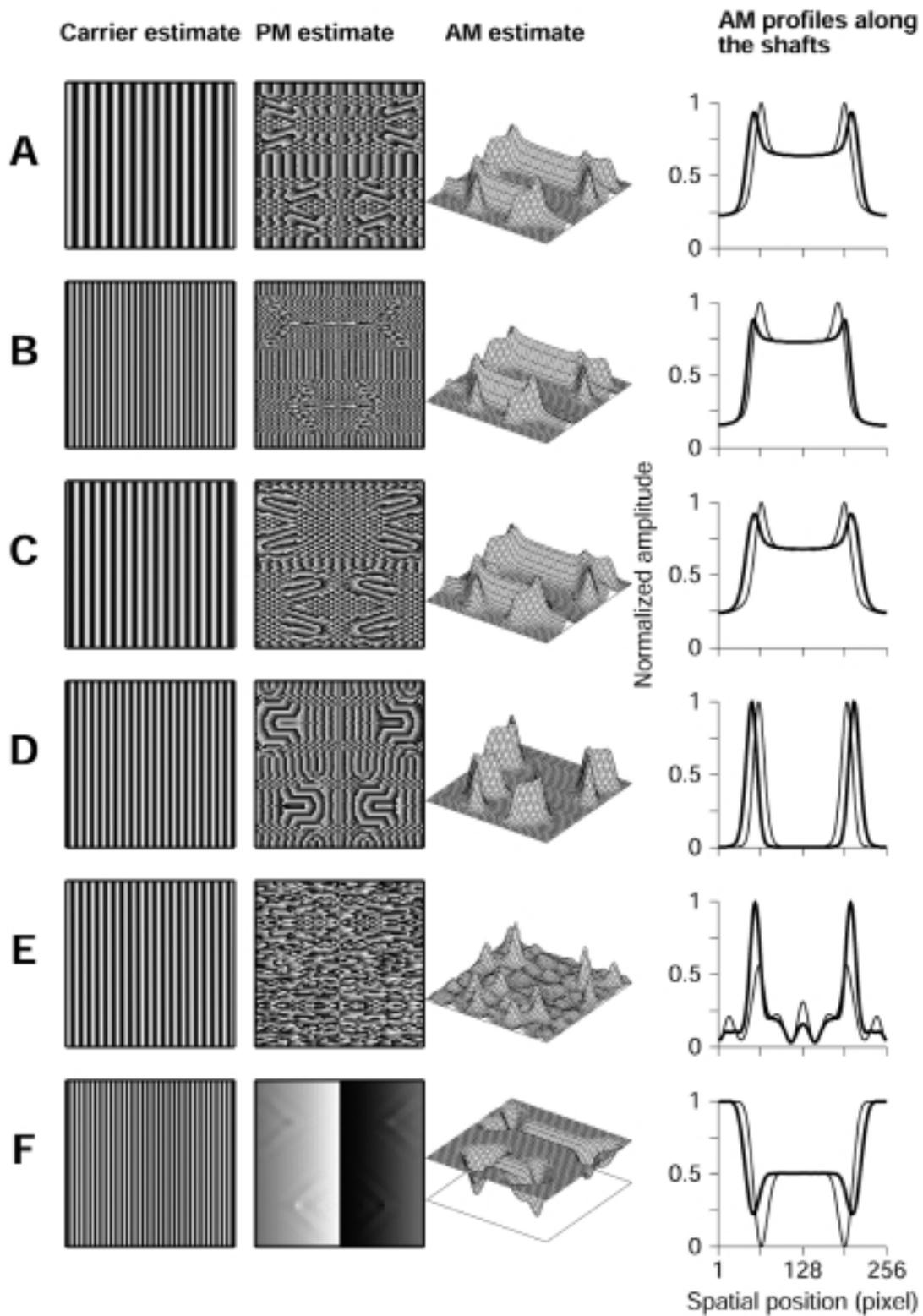


Figure 4. Demodulation transform of Müller-Lyer images by means of Daugman and Downing's (1995) AMPM algorithm and the physical presence of illusion. Input images and narrow band-pass image data (not shown) as in Figure 3 and in the same order top to bottom. Filters for spatial-frequency band selection and smoothing as in Figure 3. The pair of spatial frequencies of estimated carrier is the spectral (spatial) center of mass of right-hand Fourier half-plane. In images of PM estimates, black represents $-\pi$ rad, medium gray, 0 rad, and white, π rad. AM estimates were normalized to unity. AM profiles along the horizontal shafts have their relative amplitudes unaltered within each panel (thick line, AM profile along the upper shaft; thin line, profile along the lower shaft).

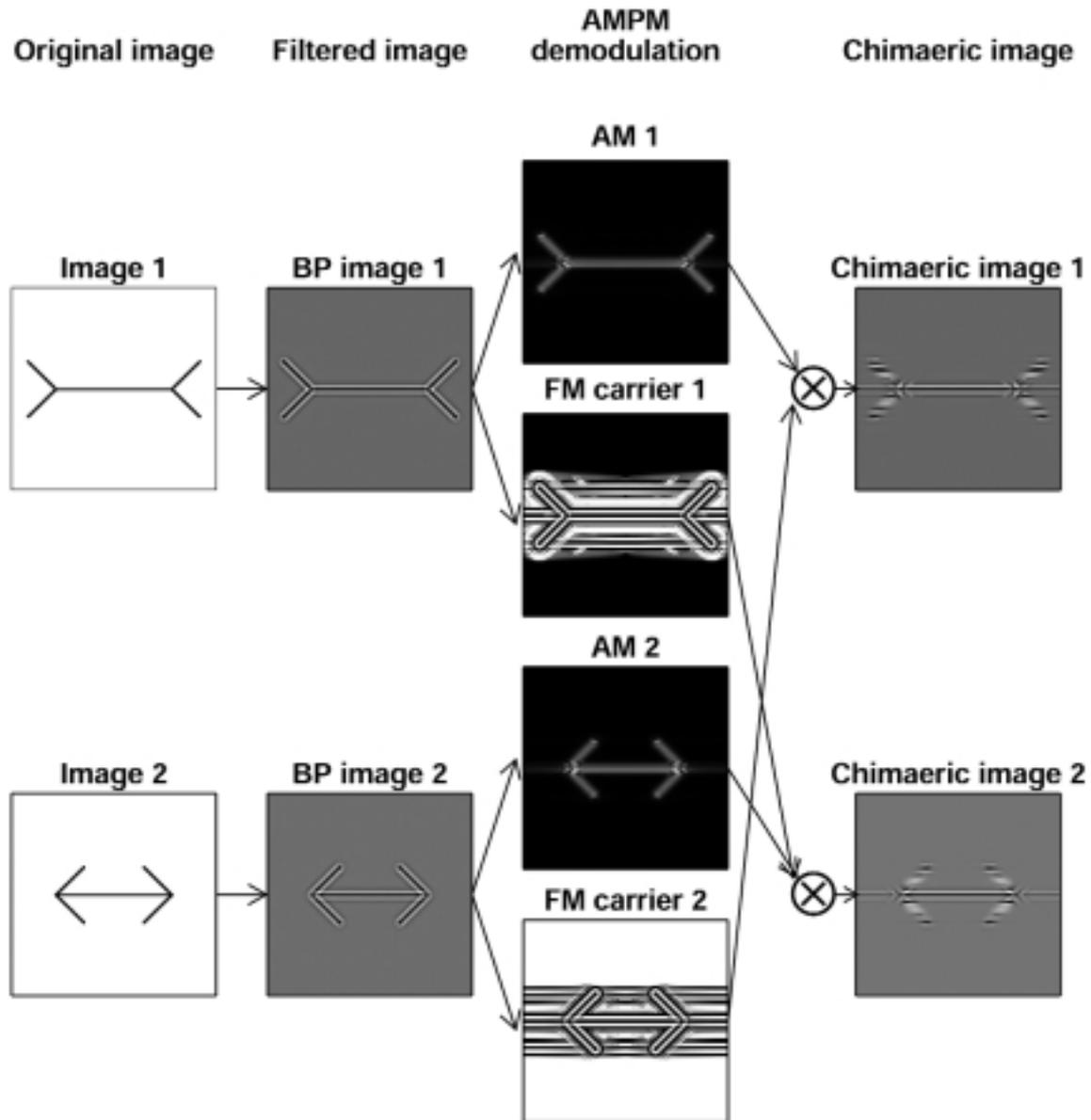


Figure 5. Visual chimaeras synthesized from the single-band AM-PM components of the two separate parts of the conventional Müller-Lyer image. Input images 1 and 2 have been band-pass filtered (the filter MTF radial profile is a cycle of a raised cosine with radial center spatial frequency of $32c/w$ and spatial frequency bandwidth of 1 octave (full-width at half height)). The AMPM algorithm was applied to each band-pass image and their AM and FM carrier estimates are shown in the third column. As depicted, each chimaeric image (fourth column) was obtained by multiplying the AM component of one image with the FM carrier of the other.

from the product of the AM component of a band-pass filtered image and the FM carrier of a different image. We used the AMPM algorithm for this purpose because reconstruction of the original the image from its AM-PM components is easier than from its DESA components. Chimaeric images in Figure 5 (fourth column) result from the combination of the AM components of one of the parts of the Müller-Lyer configuration and the FM carrier of the other. Clearly, at least for the selected spatial frequency band, the illusion remains despite the fact that the fine structure (i.e., the FM carriers) has been swapped.

A “Back-Pocket” Model of Visual Texture Segregation and the Qualitative Explanation of the Appearance of the Müller-Lyer Illusion

The two algorithms considered here provide a unique and physically reasonable solution to the demodulation problem. But it is not plausible that either algorithm could be implemented by the human visual system. Because of that, we explored the possibility that envelope estimation is carried out by the visual system in some other way. Schofield and Georgeson (2003, p. 246) have claimed that the so-called

“second-order vision” could demodulate the carrier to recover the envelope of contrast modulated periodic stimuli. Figure 6 shows a simplified working version of the linear-nonlinear-linear (LNL) model of texture segregation (or “back-pocket model”) in second-order vision (Chubb & Landy, 1991), similar to early visual mechanisms for texture perception proposed by Bergen and Adelson (1988) and Malik and Perona (1990). In a typical back-pocket model (Landy & Oruç, 2002, their Figure 1), the input image goes through a set of linear, band-pass, spatial filters (first-order channels); then, a point-wise non-linearity (full-wave rectification) is applied to the outputs; and finally, a second-order, linear, low-pass or lower spatial-frequency tuned band-pass filter is used (see the flowchart in Figure 2B). The functional architecture in the model of Landy and Oruç has oriented filters at the two linear stages; for simplicity the working model used in Figure 6 has isotropic filters. In the following, this working model is applied to the explanation of the Müller-Lyer illusion in a high-pass filtered image numerically devoid of low spatial frequencies. Given the amplitude spectrum of this input image,

first-order channels tuned to spatial frequencies lower than the cut-off spatial frequency of the stimulus are completely silent or weakly responsive (Figure 6C, top and center); on the contrary, channels tuned to higher spatial frequencies give strong responses (Figure 6C, bottom). The subsequent nonlinearity (Figure 6D) is applied to the output of each channel. Rectified outputs (Figure 6E) feed a single second-order low-pass filter (Figure 6F), but only the output of the first-order higher channel activates the second-order filter. The output of the 2D second-order-filter has spatial properties similar to the physical AM component estimated with the algorithms used above. Hence, the envelope of this image, in which the activation of low spatial-frequency first-order channels is absent, could be estimated then by a low spatial-frequency tuned band-pass second-order filter fed by the rectified outputs of active first-order channels tuned to higher spatial frequencies. In addition, the distance between the two local maxima along the location of the upper shaft in the resultant image is longer than it is for the lower one (Fig 6G, bottom), thereby explaining the perception of the illusion.

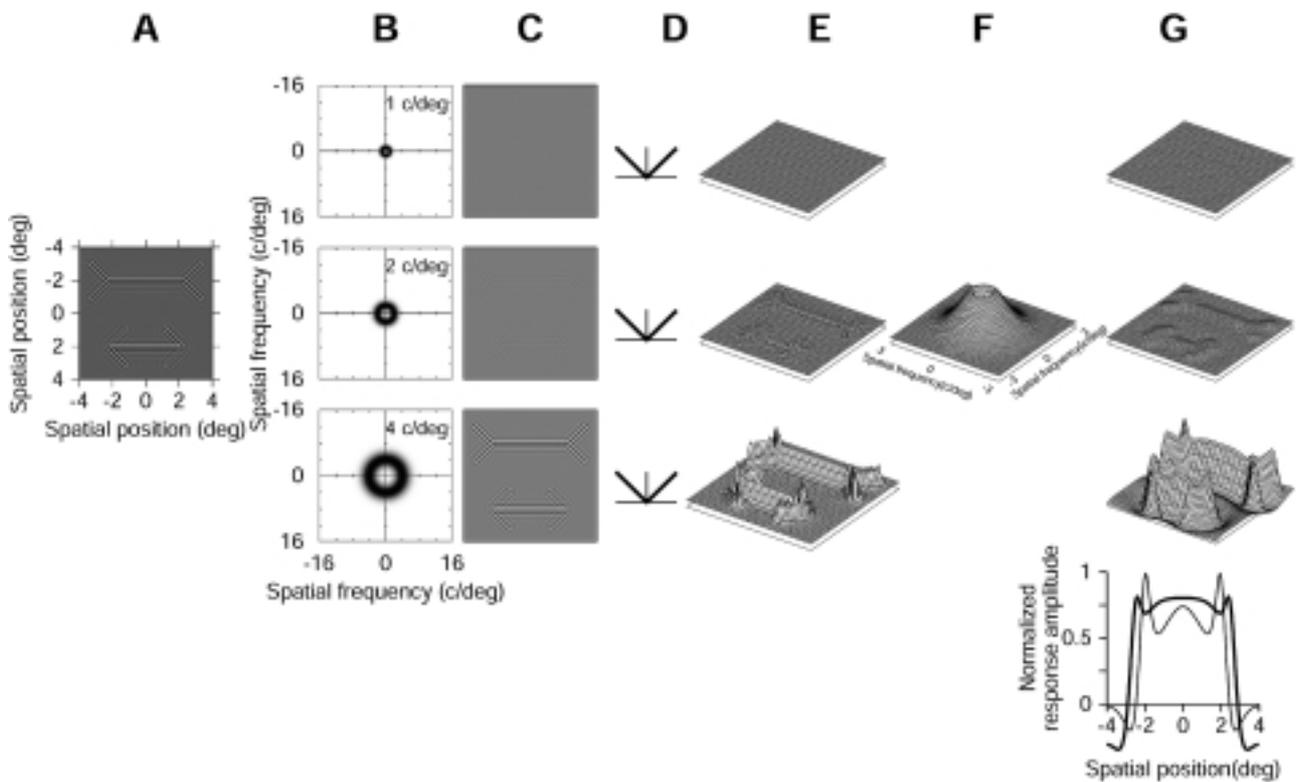


Figure 6. A simplified back-pocket model of texture segregation and the visual appearance of the Müller-Lyer illusion in an image lacking low spatial frequencies. A: The input stimulus is the high-pass filtered version in Figure 1C (it is assumed that the image subtends 8×8 deg of visual angle). B: 2D MTF support of three first-order visual channels, each consisting of the sum of oriented Gabor filters tuned to the same radial spatial frequency (1, 2, and 4 c/deg). Channel bandwidths (full-width at half height, in octaves) are calculated according of Schofield and Georgeson (2003; Equation A.6 in their Appendix A). C: First-order channel outputs (grey levels are displayed relative to the maximum amplitude across all three outputs). D: Non-linear process (full-wave rectification). E: Perspective plots of the rectified first-order channel outputs. F: Perspective plot of the 2D MTF of a second-order visual filter. The MTF is a 2D-DoG function whose radial profile fits the 1D results of Landy and Oruç (2002, their Figure 7, subject ELA); peak amplitude at a radial spatial frequency of 0.625 c/deg. Each rectified output is filtered by this single second-order visual channel. G: Perspective plot of second-order outputs; underneath, the amplitude profiles along each shaft in the second-order output at the higher frequency band (thick line, profile along the upper shaft; thin line, profile at the lower shaft).

This general process also explains the illusion in the remaining versions of the Müller-Lyer configuration used here, but dominant first-order channels tuned to other spatial frequencies are involved in those cases. Of course, if the input is a broadband image, as the conventional form of figure, all first-order visual channels would respond strongly and the illusion will also appear in each partial second-order output (results not shown here).

Discussion

This paper presents numerical results concerning the amplitude demodulation of Müller-Lyer images by means of two known AM-FM demodulation algorithms. In the following sections, we compare the performance of algorithms for AM-FM image decomposition; we examine the relationship between local amplitude and the physical presence of Müller-Lyer illusion; and, finally, we discuss the role of spatial second-order visual mechanisms in the perception of this illusion.

Comparison of DESA and the AMPM Algorithm for Image AM-FM Decomposition

It is known that high-pass filtered images can be treated as a form of 2D amplitude modulation signals (Peli, 1992); but what is more important is that this statement can also be extended to images that have been band-pass filtered with symmetric filters (Papoulis, 1962, p. 127, for 1D signals), and hence to outputs of first-order visual channels (in fact, single-band images). Decomposition of a band-pass image lacking explicit contrast modulation in its AM and FM (or PM) components seems useful and natural because the AM component carries information about local energy in the image, such as local image contrast, whereas the FM (or PM) component contains information about the local spatial structure of the image, such as image texture (Havlicek & Bovik, 2000; Felsberg & Sommer, 2001). Both of the algorithms used here fulfill the requirements of complete image reconstruction from the estimated components (almost exact, in case of the AMPM algorithm) and the AM estimates have physical meaning. For the same Müller-Lyer image, there are slight differences in AM estimates that are caused by the different procedures used by the algorithms: DESA uses a *local* operator over the *real* or *complex* values of a given signal, whereas the AMPM algorithm is a *global* operator that uses all *complex values* in the half Fourier plane in one *arbitrary direction* to obtain point-wise estimates. AM components estimated by DESA were, in all cases, spatially localized and they had compact spatial support (Figure 3, third column). On the contrary, AM components estimated by the AMPM algorithm were, in some cases, elongated horizontally (Figure 4, third column). This is because the AMPM algorithm is a directional (or anisotropic) algorithm and it suffers from the same complications as

directional 2D Hilbert transforms (i.e., the 1D Hilbert transform with respect to an arbitrary preference direction, which might be one of the axes of the coordinate system). (For a definition see Bülow & Sommer, 1999). The quality of anisotropic algorithms for envelope estimation depends on the orientation of the local features in the image (Bülow & Sommer; Felsberg & Sommer, 2001). When local features are not parallel to the selected direction for the half Fourier plane (the right-hand Fourier plane in our implementation of the AMPM algorithm), the envelope is estimated well: this is the case of the calibration test image in which the orientation of main local features is orthogonal to selected direction, and it is also the case of the Brentano-Gregory figure with only oblique strokes (Figure 4D, third column). But envelope anisotropic estimation fails when local features exist that are parallel to the selected direction: This is the case of Müller-Lyer configurations with shafts oriented along the x axis (Figure 4, A, B, C, third column). Note in Figure 5 (third column) that parallel horizontal striations in the AM and FM carrier estimates exist when both components are displayed as images (compare these spurious features with horizontal lines in the AM estimate for a circular fringe obtained with the directional 2D Hilbert transform by Bülow & Sommer in their Figure 5, fourth image from the left). However, problems originated by the use of anisotropic demodulation procedures can be resolved by means of application of the Riesz transform (the multidimensional isotropic generalization of the Hilbert transform) and monogenic signal (Felsberg & Sommer) to AM-PM demodulation of illusion images (Sierra-Vázquez & Serrano-Pedraza, 2004, 2006).

Local Amplitude (or Envelope), Global Spatial Structure, and the Müller-Lyer Illusion

The Müller-Lyer illusion is perceived in images with a variety of spatial frequency content. The appearance of the illusion in broadband images (e.g., the conventional Müller-Lyer version in Figure 1A, the Brentano-Gregory configuration in Figure 1F, or the Skottun flat spectrum image in Figure E) does not challenge Ginsburg's explanation: In fact, contrary to the claims of Gregory (1990) and Skottun (2000), low-pass filtering of these images physically creates the illusion. The point at issue is the perception of the illusion in images devoid of visible low-spatial frequencies (e.g., the narrow band-pass image in Figure 1D or strictly high-pass version in 1C), images which bypass the problem of residual low-spatial frequencies addressed by García-Pérez (1991) for a balanced-square figure (Figure 1B). Narrow band-pass filtering for the selection of a single frequency band (taken as a measure of spatial scale) automatically introduces a spatial modulation of contrast into all the images used here. Subsequent single-band amplitude demodulation shows that the illusion *is physically present in the spatial envelope of the filtered image obtained as a result of the selection of*

a single spatial frequency band (or scale) and not exclusively in the low-spatial frequency range (or coarse scales), as was misinterpreted from digital or optical blurring of illusory configurations (Coren & Girgus, 1978, their Figure 3; Ginsburg, 1978, his Figure 50 a and b). Furthermore, multiband demodulation (results not presented here) shows that the Müller-Lyer illusion is physically present in the AM component of each single band. Of course, synthetic images (as some of Dalí's paintings) could be generated in which each spatial scale carries a different local amplitude and local phase, but this is not the case of the (broadband) conventional Müller-Lyer illusion image, all of whose spatial frequency bands contain similar global structural information.

The AM component or envelope estimated by means of DESA and the AMPM algorithm presents low-pass spatial characteristics and, thus, it contains information about the global structure of the image. A widespread belief in the field of spatial vision is that the whole of an image is associated with the lower spatial frequencies (or coarse scales) and its details with the higher ones (or fine scales). However, demodulation algorithms undermine this belief. Demodulation shows that every single spatial scale (alternatively, every single spatial frequency band) contains information both about the whole and the fine spatial structure of the image: single-band AM components carry information about the global spatial structure of an image and FM (or PM) components carry information about its fine spatial structure. As chimaeric images of the Müller-Lyer illusion demonstrate, it is precisely the envelope that carries the global spatial structure of the image at the same time as the illusion.

The decisive contribution of the envelope to the appearance of the illusion is not in contradiction with Skottun's (2000) statement that the appearance of the illusion is mainly associated with the Fourier phase spectrum rather than with its Fourier amplitude spectrum. Skottun's (2000) demonstration merely confirms the known fact that spatial structure is carried by the Fourier phase spectrum (under some conditions; see Tadmor & Tolhurst, 1993) and, thus, the appearance of hybrid images is like that of the image that contributed the phase spectrum (Oppenheim & Lim, 1981; Piotrowski & Campbell, 1982). But the appearance of the illusion still remains unexplained: In hybrid images, the lengths of the shafts are physically equal (Skottun's Figure 3). However, AM-FM demodulation of these images would again show that their local amplitudes present spatial distortions in the direction of perceived illusion.

Second-Order Spatial Vision and the Perception of the Illusion

It is known that the human visual system can demodulate the amplitude component of so-called contrast-modulated (or second-order) stimuli and recover their

periodic envelope (Schofield & Georgerson, 2003). To explain the appearance of the Müller-Lyer illusion, we suggest that the human visual system can also demodulate complex stimuli and recover the non-periodic envelope where the information about global structure is found. The specific way in which the visual system does this could be discussed. The front end of the human visual system consists of a set of band-pass first-order channels that, at the same time as it selects different spatial-frequency bands, introduces contrast modulation into the filtered stimuli. In addition, as an alternative to Ginsburg's (1978) explanation, Carlson et al. (1984) suggested that there could be a non-linearity in the visual system before the low-pass frequency channel, introducing low spatial frequencies in the input image. The current back-pocket model puts together the linear front-end, the non-linearity, and the low-pass or low spatial frequency tuned band-pass visual filter, which is applied to the result of the non-linear process (whichever it is), and not directly to the input image, as Ginsburg proposed. The sequence filter-rectify-filter used in the model of Figure 6 describes the simplest possible second-order or non-Fourier mechanism for visual amplitude demodulation. The aim of the example in Figure 6 is not to validate a specific model of second-order spatial vision, but to illustrate that the proposed visual architecture can demodulate the carrier (as a numerical demodulation algorithm does) and recover the envelope that carries the illusion. More complex models propose a variety of mappings of first-order channels onto second-order channels and embody additional mechanisms while they maintain the intermediate non-linear rectification stage (see Schofield, 2000). We used intermediate full-wave rectification and low-pass filtering according to more extended models, but this is not the only demodulation procedure that could be implemented by the human visual system. Other popular models, based on visual physiology, embody a set of directional "analytic" filters defined over the Fourier half-plane as daisy petal shapes (or oriented Gabor filters in the spatial domain): The local amplitude and local phase of the image are, respectively, the modulus and argument of the corresponding complex signals of their outputs (Bovik, Clark, & Geisler, 1990; Daugman & Downing, 1995; Havlicek & Bovik, 2000).

To reconcile his own numerical results with the experimental results of Carrasco et al. (1986), Skottun (2000) suggested that adaptation to a sinusoidal grating "may have had its effect on the illusion indirectly via the image phase spectra" (p. 208). We suggest that there is a simpler and more comprehensive explanation of these results in terms of spatial amplitude demodulation processes carried out by second-order visual mechanisms. In fact, the explanation of adaptation results of Carrasco et al. is theoretically similar to the explanation of illusion appearance in high-pass filtered stimuli. In a well-behaved broadband stimulus (such as the conventional Müller-Lyer

configuration), energy concentrates in the low-spatial frequency range (Figure 1A). If the illusion were exclusively carried by the lower spatial-frequency components in the Fourier amplitude spectrum and were due to subsequent visual low-pass filtering, the fatigue of these channels by adaptation to a sinusoidal grating of relatively low spatial frequency (2 c/deg) should result in the complete disappearance of the illusion. But this is not the case in the results reported by Carrasco et al.: The strength of the illusion (measured as the ratio between the physical lengths of the shafts when perceived as equal) decreases from 1.32 to 1.26, but the illusion does not disappear. If the visual stimulus is a broadband image (as is the case), adaptation to a sinusoidal grating of low spatial frequency effectively reduces or completely eliminates the activity of lower spatial frequency tuned visual channels, but it has no effect on active higher spatial frequency first-order channels that feed amplitude demodulation (second-order) mechanisms (as shown in Figure 6). In consequence, after adaptation, the illusion, even though reduced, persists.

Admittedly, as pointed out by Coren and Girgus (1978), there is not a single cause of geometrical illusions. But it is not reasonable to suppose, as it is done in some works, ad hoc processes for each geometrical illusion and a homunculus that chooses the appropriate mechanisms so that the illusion appears. Amplitude demodulation processes explain visual phenomena related to the perception of global patterns such as preattentive texture segregation (Schofield, 2000), spatiotemporal beats, illusory contours, and grouping (Daugman & Downing, 1995) besides geometrical illusions (Sierra-Vázquez & Serrano-Pedraza, 2006), suggesting a unified framework for a parsimonious explanation of these phenomena of spatial vision.

Conclusions

To conclude:

1. The Müller-Lyer illusion is physically carried in the AM component of single band-pass filtered images in which the illusion is perceived, including images lacking low spatial frequencies.
2. The Müller-Lyer illusion could appear as a result of image envelope recovery by the visual system. The proposed back-pocket model for preattentive visual texture segregation could carry out the estimation of the AM component of stimuli, thus explaining the perception of the illusion in all the cases studied here.
3. Therefore, the illusion in conventional and derived Müller-Lyer configurations is not a separate problem in spatial vision but arises as a natural consequence of second-order processing, visual mechanisms being identically the same both in normal preattentive visual texture perception and in illusion.

References

- Bergen, J.R., & Adelson, E.H. (1988). Early vision and texture perception. *Nature*, *333*, 363-364.
- Bovik, A.C., Clark, M., & Geisler, W.S. (1990). Multichannel texture analysis using localized spatial filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12*, 55-73.
- Bovik, A.C., Gopal, N., Emmoth, T., & Restrepo (Palacios), A. (1992). Localized measurement of emergent image frequencies by Gabor wavelets. *IEEE Transactions on Information Theory*, *38*, 691-712.
- Boring, E.G. (1942). *Sensation and perception in the history of experimental psychology*. New York: Appleton-Century-Crofts.
- Bracewell, R.N. (1978). *The Fourier transform and its applications* (2nd ed.). New York: McGraw-Hill.
- Brentano, F. (1892). Über ein optisches Paradoxen. *Zeitschrift für Psychologie*, *3*, 349-358.
- Bülöw, T., & Sommer, G. (1999). A novel approach to the 2D analytic signal. In F. Solina & A. Leonardis (Eds.), *8th Conference on Computer Analysis of Images and Patterns, Ljubljana* (Vol. 1689 of LNCS, pp. 25-32). Berlin: Springer-Verlag.
- Carlson, C.R., Moeller, J.R., & Anderson, C.H. (1984). Visual illusions without low spatial frequencies. *Vision Research*, *24*, 1407-1413.
- Carrasco, M., Figueroa, J.G., & Willen, J.D. (1986). A test of the spatial-frequency explanation of the Müller-Lyer illusion. *Perception*, *15*, 553-562.
- Coren, S., & Girgus, J.S. (1978). Visual illusions. In R. Held, H.W. Leibowitz, & H.-L. Teuber (Eds.), *Handbook of sensory physiology. Vol VIII. Perception* (pp. 551-568). Berlin: Springer-Verlag.
- Chubb, C., & Landy, M.S. (1991). Orthogonal distribution analysis: A new approach to the study of texture perception. In M.S. Landy & J.A. Movshon (Eds.), *Computational models of visual processing* (pp. 291-301). Cambridge, MA: MIT Press.
- Daugman, J.G., & Downing, C.J. (1993). Demodulation: A new approach to texture vision. In L.H. Harris & M. Jenkin (Eds.), *Spatial vision in human and robots* (pp. 63-87). Cambridge, UK: Cambridge University Press.
- Daugman, J.G., & Downing, C.J. (1995). Demodulation, predictive coding, and spatial vision. *Journal of the Optical Society of America A*, *4*, 641-660.
- Felsberg, M., & Sommer, G. (2001). The monogenic signal. *IEEE Transactions on Signal Processing*, *49*, 3136-3144.
- García-Pérez, M.A. (1991). Visual phenomena without low spatial frequencies: A closer look. *Vision Research*, *31*, 1647-1653.
- Ginsburg, A.P. (1978). *Visual information processing based on spatial filters constrained by biological data*. Technical Report AMRL-TR-78-129. Vol I & II. Aerospace Medical Research Laboratory. Dayton, OH: Wright Patterson Air Force Base.
- Ginsburg, A.P. (1986). Spatial filtering and visual form perception. In K. Boff, L. Kauffman, & J.P. Thomas (Eds.), *Handbook of perception and human performance. Vol II. Cognitive processes and performance* (pp. 34.1-34.41). New York: Wiley.
- Goldstein, E.B. (1984). *Sensation and perception*, (2nd ed.). Belmont, CA: Wadsworth.

- Gregory, R. (1990). *Eye and brain. The psychology of seeing* (4th ed.). Oxford, UK: Oxford University Press.
- Hartmann, W.M. (1998). *Signals, sound, and sensation*. New York: Springer-Verlag.
- Havlicek, J.P. (1996). AM-FM image models. http://www.ece.utexas.edu/~bevans/courses/ee381k/lectures/am_fm_models/AM_FM_Image_Models.pdf.
- Havlicek, J.P., & Bovik, A.C. (2000). Image modulation models.. In A.C. Bovick (Ed.), *Handbook of image and video processing* (pp. 313-324). New York: Academic Press.
- Heymans, G. (1896). Quantitative Untersuchungen über das "optische Paradoxon". *Zeitschrift für Psychologie*, 9, 221-255.
- Kaiser, J.F. (1990). On a simple algorithm to calculate the energy of a signal. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, April, 381-384.
- Landy, M. S., & Oruç, I. (2002). Properties of second-order spatial frequency channels. *Vision Research*, 42, 2311-2329.
- Loughlin, P.J., & Tacer, B. (1996). On the amplitude- and frequency-modulation decomposition of signals. *Journal of Acoustical Society of America*, 100, 1594-1601.
- Malik, J., & Perona, P. (1990). Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America A*, 7, 923-932.
- Maragos, P., & Bovik, A.C. (1995). Image demodulation using multidimensional energy separation. *Journal of the Optical Society of America A*, 12, 1867-1876.
- Maragos, P., Kaiser, J.F., & Quatieri, T.F. (1993). On amplitude and frequency demodulation using energy operators. *IEEE Transactions on Signal Processing*, 41, 1532-1550.
- Morrone, M.C., & Burr, D.C. (1988). Feature detection in human vision: A phase-dependent energy model. *Proceedings of the Royal Society, London, B*, 235, 221-245.
- Müller-Lyer, F.C. (1889). Optische Urtheilstäuschungen. *Archiv für Anatomie und Physiologie, Physiologische Abteilung 2* (Suppl.), 263-270.
- Müller-Lyer, F.C. (1896). Zur Lehre von den optischen Täuschungen. Über Kontrast und Konfluxion. *Zeitschrift für Psychologie*, 9, 1-16.
- NAG (1991). *NAG Fortran Library manual-Mark 15*. Oxford: Numerical Algorithm Group.
- Oppenheim, A.V., & Lim, J.S. (1981). The importance of phase in signals. *Proceedings of the IEEE*, 69, 529-541.
- Papoulis, A. (1962). *The Fourier integral and its applications*. New York: McGraw-Hill.
- Peli, E. (1992). Perception and interpretation of high-pass filtered images. *Optical Engineering*, 31, 74-81.
- Piotrowski, L.N., & Campbell, F.W. (1982). A demonstration of the visual importance and flexibility of spatial-frequency amplitude and phase. *Perception*, 11, 337-356.
- Robinson, J.O. (1972). *The psychology of visual illusion*. London: Hutchinson University Library.
- Schofield, A.J. (2000). What does second order-vision see in an image? *Perception*, 29, 1071-1086.
- Schofield, A.J., & Georgeson, M.A. (2003). Sensitivity to contrast modulation: The spatial frequency dependence of second-order vision. *Vision Research*, 43, 243-259.
- Sekuler, R., & Blake, R. (1994). *Perception* (3rd ed). New York: McGraw-Hill.
- Skottun, B.C. (2000). Amplitude and phase in the Müller-Lyer illusion. *Perception*, 29, 201-209.
- Sierra-Vázquez, V., & García-Pérez, M.A. (1995). Psychophysical 1-D wavelet analysis and the appearance of visual contrast illusions. *IEEE Transactions on Systems, Man, and Cybernetics*, 25, 1424-1433.
- Sierra-Vázquez, V., & Serrano-Pedraza, I. (2003, August). Anisotropic amplitude demodulation of Müller-Lyer illusion. *34th European Mathematical Psychology Group Meeting* (abstract). Madrid,
- Sierra-Vázquez, V., & Serrano-Pedraza, I. (2004). Riesz transforms for the isotropic envelope estimation of Müller-Lyer illusion images. *Perception* (Suppl.), 33, 81 (abstract).
- Sierra-Vázquez, V., & Serrano-Pedraza, I. (2006, August). Application of Riesz transforms to AM-PM decomposition of geometrical optical illusions images. *International Congress of Mathematicians*, Madrid. *Abstracts*, 178.
- Smith, Z.M., Delgutte, B., & Oxenham, A.J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, 416, 87-90.
- Tadmor, Y., & Tolhurst, D.J. (1993). Both the phase and the amplitude spectrum may determine the appearance of natural images. *Vision Research*, 33, 141-145.
- Vakman, D.E. (1972). On the definition of concepts of amplitude, phase, and instantaneous frequency of a signal. *Radio Engineering and Electronics Physics*, 17, 754-759.

Received January 25, 2006

Revision received September 8, 2006

Accepted October 21, 2006

Appendix

AM-FM Demodulation Procedures and Performance of the Algorithms

Test Image

The test image (Figure A1) is a synthetic 2D AM-FM signal given by,

$$I(m,n) = I_0[1 + s(m,n)], \quad m = 1, \dots, M, \quad n = 1, \dots, N, \quad (\text{A.1})$$

with,

$$s(m,n) = k[1 + f_{LP}(m,n)] \cos\left[2\pi(u_0 \frac{n}{N} + v_0 \frac{n}{N}) + \sin(2\pi u_1 \frac{n}{N} + \pi/2)\right], \quad (\text{A.2})$$

where the contrast modulating function, f_{LP} , $-1 \leq f_{LP}(i,j) \leq 1$, is a low-pass filtered image of letter C and the parameter values are: $I_0 = 128$, $M = 128$, $N = 128$, $k = 0.5$, $u_0 = 10$ c/w, $v_0 = 0$ c/w, $u_1 = 2$ c/w.

For this signal, its envelope, or AM component, is,

$$|a(m,n)| \equiv 2A(m,n) = k[1 + f_{LP}(m,n)],$$

and the FM carrier (c in Equation 1) is

$$c(m,n) = \cos\left[2\pi u_0 \frac{n}{N} + \sin(2\pi u_1 \frac{n}{N} + \pi/2)\right].$$

According to the discrete version of Equation 2, the vertical and horizontal instantaneous spatial frequencies of test image are, respectively,

$$\Omega_1(m,n) = 0$$

and

$$\Omega_2(m,n) = \frac{1}{N} \left[2\pi u_0 + 2\pi u_1 \cos(2\pi u_1 \frac{n}{N} + \pi/2) \right]$$

The AM component and FM functions are depicted as surfaces in Figure A1 (column A).

Alternatively, the sinusoidal carrier (C in Equation 3) is $C(m,n) = \cos(2\pi u_0 n/N)$, and the PM component is $\varphi(m,n) = \sin(2\pi u_1 n/N + \pi/2)$. These AM and PM components are depicted as images in Figure A2 (column A). Due to the narrow-band nature of the test image (as its amplitude spectrum in Figure A1 indicates), no band-pass prefiltering was necessary before the application of the algorithms.

Performance of the Algorithms

The main requirements of the algorithms are that the resultant AM estimate should be the envelope of the original test signal and that they should allow recovery of the original FM function (DESA) or carrier and PM component (AMPMP algorithm). The capability of the algorithms to track these components was measured for the AM component in DESA and for each component in the AMPMP algorithm. Measures were: (a) the extremes of range of the component estimate; (b) the normalized mean square error, $\text{NMSE} = e_{\text{rms}}^2 / \sigma^2$, where e_{rms}^2 is the mean square error for the component estimate, and σ^2 is the variance of the true values of the component; and (c) the extremes of range of the absolute errors (defined as absolute pointwise differences between the true and the estimated values for each component). Estimation of the pair of spatial frequencies for the sinusoidal carrier and image reconstruction apply only to the AMPMP algorithm.

The DESA

Following the indications of Maragos and Bovik (1995, pp. 1871-1872) for the DESA, we compute the amplitude envelope, $|a(m,n)|$, the ($[0, \pi/2]$ -bounded) vertical instantaneous frequency, $|\Omega_1(m,n)|$, and the ($[0, \pi/2]$ -bounded) horizontal instantaneous frequency, $|\Omega_2(m,n)|$, according to their Equations 60, 58, and 59, respectively, after application of a 2D discrete-space energy operator (based on a discrete-space counterpart of the TKEO). Numerical results are shown in Figure A1. Raw estimates are shown in column B. When applied to our test image, DESA reveals some representative problems. One of them is the presence of spikes, which are outlier estimates caused by numerical singularities. In addition, a small number of negative amplitude values occur which are absurd (see in Figure A1, column B, the AM component). Paradoxical FM estimates (spikes in column B, second and third rows) are mainly due to the local character of TKEO. To be fair, and because the purpose of this calibration is not the discussion of the pros and cons of this algorithm about FM function recovery, border and edge effects in instantaneous spatial frequency (noted by Havlicek & Bovik, 2000) were simply removed. Raw results were smoothed with a low-pass post-filter to stabilize the solutions, eliminate these outlier estimates, and remove negative amplitudes (Havlicek, 1996). Smoothed results are shown in column C, pointwise absolute errors (original quantities minus smoothed estimates) in column D, and profiles in column E. Low-pass filtering removes spikes in AM raw estimates but it does not distort the estimate (see the AM component profile in column E). In addition, filtering attenuates negative amplitude values, but, unfortunately, it does not remove all of them. Because negative values do not make any physical sense, in our computations, we set them all equal to zero, as Maragos and Bovik did. Despite these negatives values, raw and smoothed estimation of the AM component is almost exact (NMSEs are 0.013 and 0.008, respectively). On the contrary, instantaneous spatial frequency estimates suffer from the abrupt spatial transitions in the test image: Estimates of vertical frequency modulation, even though flattened, are wrong, whereas horizontal ones are correct outside the space between horizontal parts of letter C (see the profiles in column E).

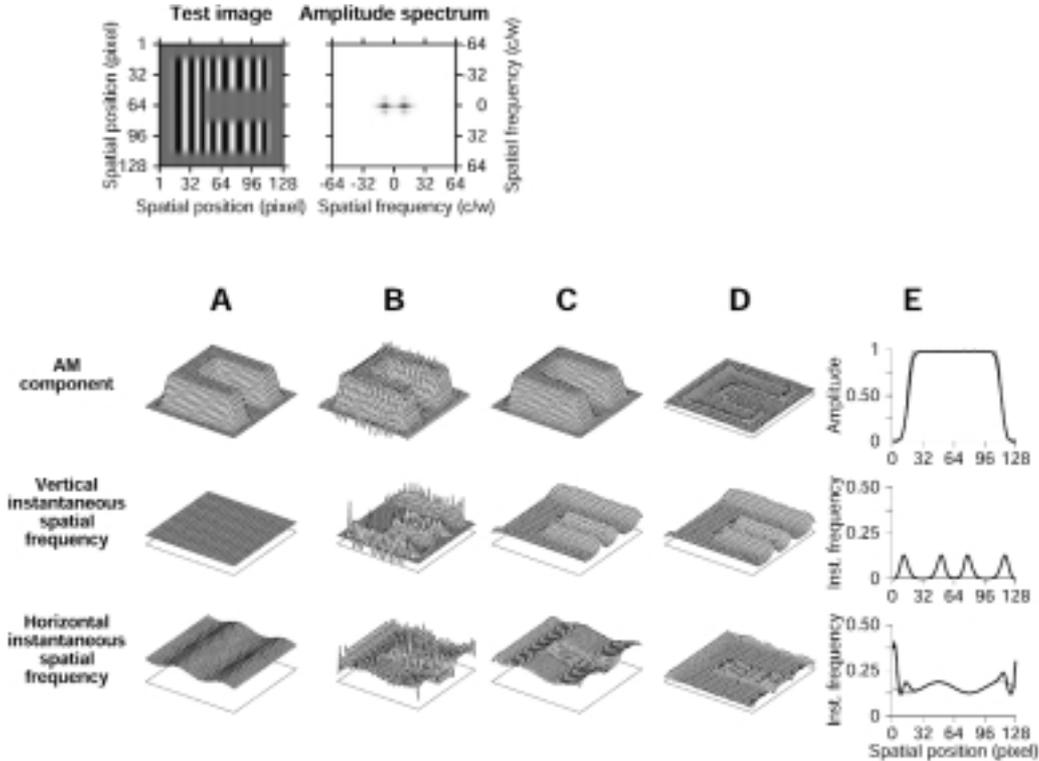


Figure A1. Application of the 2D TKEO and the DESA to a synthetic 2D AM-FM test image. The upper left test image is the contrast modulated FM carrier defined in Equations A.1 and A.2; its amplitude spectrum is shown on the right. A: Perspective plot of the respective original modulation components of the test image. Values of vertical and horizontal instantaneous spatial frequencies are divided by π . B: Perspective plot of the raw estimates of image modulation components obtained applying the DESA to the test image. Spikes are outlier estimates. C: Results of smoothing the raw estimates. The radial profile of the MTF of the isotropic low-pass smoothing filter is a Gaussian function with a standard deviation of 8 c/w. D: Perspective plot of absolute errors defined as the pointwise difference between the original and the smoothed estimate values for each component. E: Profile of the modulation components. Thin line, original component; thick line, smoothed estimate. Profiles of original and smoothed AM component and original and smoothed horizontal instantaneous spatial frequency estimates were along the row $m = 32$. Profile of original and smoothed vertical instantaneous spatial frequency were along the column $n = 64$.

The AMPM Algorithm

Following the indications of Daugman and Downing (1993, pp. 65-69) for the AMPM algorithm, we compute the pairs (u_c, v_c) , $C(m, n)$, $A(m, n)$, and $\varphi(m, n)$ according to their Equations 3 to 7, and 12 and 13. Numerical results are shown in Figure A2. Raw results and image reconstruction from raw component estimates are shown in column B. Notice that, in this algorithm, PM estimation depends on carrier spatial frequencies (in our computations, the spectral center-of-mass, but the actual values of carrier frequencies are unimportant); on the contrary, the AM component estimation depends only on the signal and not on the choice of carrier spatial frequencies (Daugman & Downing, 1995, their Equation 14). The raw AM estimate was smoothed with the same low-pass filter used in DESA (result shown in column C). Absolute pointwise errors are shown in column D, and profiles in column E. Again, raw and smoothed estimations of the AM component are almost exact (NMSEs are 0.0057 and 0.0106, respectively) and so is the estimation of the carrier, including the pair of spatial frequencies. Absurd absolute error range in local phase estimate is mainly due to phase wrapping, as are values in borders and inside the uniform region within letter C; outside these areas, the AMPM algorithm gives realistic (but not exact) estimates of local phase (see PM component profiles in column E). More important, reconstruction (without normalization) of the input image from its AM-PM decomposition is almost exact (NMSE equal to 0.0107).

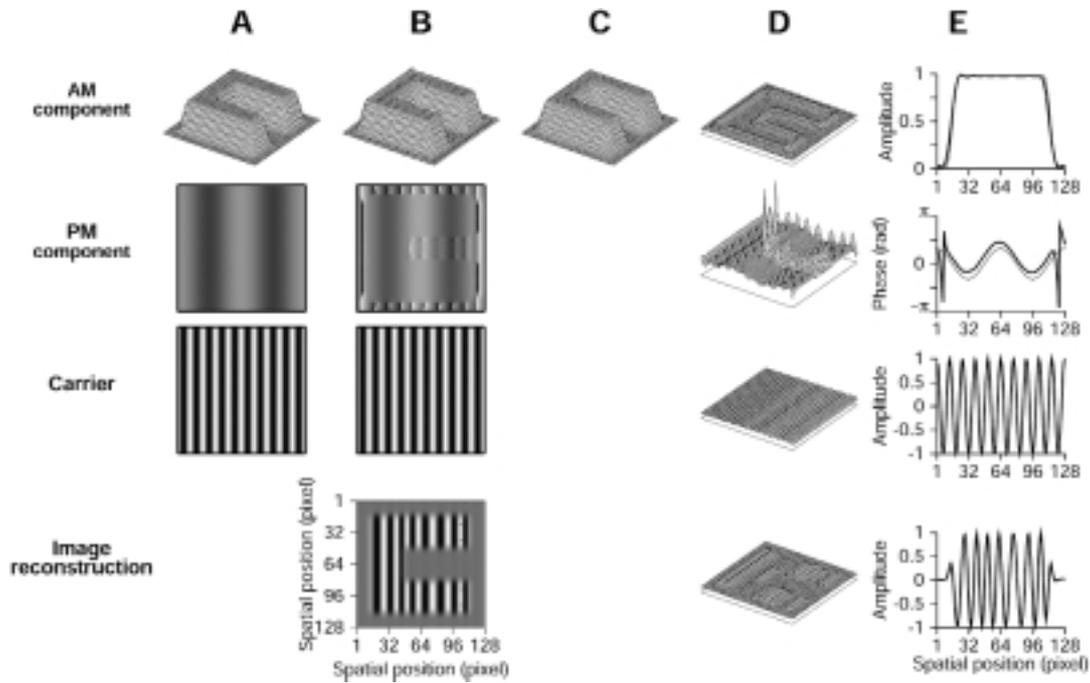


Figure A2. Application of the AMPM algorithm to the test image in the upper left of Figure A1. A: Perspective plot of the original AM component and images of the PM component (black, $-\pi$, medium grey, 0, and white, π) and the sinusoidal carrier. B: Perspective plot of raw AM estimates and images of raw PM and carrier estimates. The bottom panel shows test image reconstruction from raw estimated components. C: Perspective plot of the result of smoothing the raw AM estimation (type and parameters of the smoothing filter as in Figure A1). D: Perspective plot of absolute errors defined as the pointwise difference between the original and raw estimate values for each component. E: Profile along the 32nd row in the images of original and estimated components. Thin line, original component; thick line, estimated component.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.