

Syntactical Visual Analysis and Reconstruction

du Buf, J.M.H. and Fischer, S.*

Signal Processing Laboratory,
Swiss Federal Institute of Technology, CH-1015 Lausanne, Switzerland

Abstract: Utilising the complex Gabor code as an equivalent of the cortical image representation, this code must not only allow to extract syntactical information like differently shaped lines and edges but also to synthesise a replica of the visual field. Instead of utilising the Gabor code directly in the image synthesis, as is common in digital image coding, we study the possibility of a syntactical coding and reconstruction based on multiscale lines and edges, at the moment in the 1D case.

Introduction

Complex Gabor functions are a mathematical abstraction of the even- and odd-symmetric oriented receptive fields of cortical simple cells. These cells (Gabor filters) provide the input code for line and step-edge (event) detectors. Isolated events can be accurately detected by local maxima in the modulus of the Gabor responses, and the phase value there gives the event type: a positive line is characterised by a phase value of 0, a negative line by $\pm\pi$, whereas positive and negative step edges show phase values of $\pm\pi/2$ [2,4]. Generalised events, that is sharp events blurred by a Gaussian lowpass filter, can also be characterised by a scale estimate.

Recent work has shown that the maximum modulus cannot give a precise position estimate when two events are close to or cross each other, and that local phase values may give more reliable position estimates [2]. Furthermore, ramp edges are also characterised by phase values of 0 and $\pm\pi$, which means that a detection scheme based on maximum modulus and phase values cannot discriminate between lines and ramp edges. This problem can be solved by adopting a syntactical image reconstruction which is based on summing event templates consisting of Gaussian lines and error function-shaped edges with the appropriate scales, and this leads to a better explanation of Mach bands as perceived at ramp edges [3].

Summarising, what we are looking for is a model of visual perception in which a precise syntactical analysis and image representation are coupled to a visual reconstruction theory.

*Supported by the Swiss National Science Foundation, grant number 21-33641.92

Gabor Scale Space

We take a 1D Gabor basis function

$$\begin{aligned}\mathcal{G}(x|\sigma, \omega_0) &= \exp\{-x^2/2\sigma^2 + i\omega_0 x\} \\ &= \exp(-x^2/2\sigma^2) \cdot \{\cos \omega_0 x + i \sin \omega_0 x\},\end{aligned}$$

where σ is the spread of the Gaussian envelope, ω_0 is the frequency, and $i = \sqrt{-1}$. This function is linearly scaled by s , giving the set $\mathcal{G}_s(x|\sigma s, \omega_0/s)$. At the moment we assume continuity, i.e. we obtain filter responses continuous in both x and s . We also assume that the modulation transfer function is normalised, i.e. $\mathcal{G}_s(x|\sigma s, \omega_0/s)/s\sigma\sqrt{2\pi}$, in order to obtain event-scale selectivity. This requires that $\omega_0 = \sqrt{1 + 1/\sigma^2}$ [see 2]. The input signal is correlated with the set \mathcal{G}_s and event-scale selectivity leads to precise scale estimates when the events are Gaussian lines or error function step edges. However, other signals like a trapezoidal edge lead to a scale estimation error: what we obtain then is an equivalent scale. In this case the equivalent error function-shaped edge must be refined by taking into account the (Gaussian) lines detected at the two ramp edges. Furthermore, ideal (sharp) events lead to many scales at which events are detected, and many of these scales must be used to reconstruct sharp structures (e.g. a triangular line with a sharp maximum).

Multiscale Analysis and Synthesis

An example of the analysis/synthesis system is shown in Figs. 1-3. The input signal consists of a trapezoidal line (Fig. 1) which evokes line responses at many scales (bold lines in Fig. 2) which are unstable at coarse scales. This is caused by interference effects. Each point on a bold line represents a line because of phase values of 0 and $\pm\pi$, each with an equivalent line scale (i.e. an unlimited Gaussian).

From all information we select only the scale intervals with reliable information. This analysis is similar to that applied in the zero-crossing scale space [1], and may be based on different stability criteria. Possible criteria are (1) the stability of the maximum modulus position, (2) the stability of the maximum phase congruence position, i.e. the minimum phase variance over neighbouring scales, (3) the zero crossings of the real and

imaginary parts of the responses and (4) the stability of response amplitudes. Our experiments have shown that the first three methods give comparable results. Figure 3 shows the resulting scale space when the maximum modulus position is used as a detection criterion. The events are reconstructed by assuming Gaussian lines, and in the general case also error function-shaped edges, with the appropriate positions, amplitudes, scales, and polarities. The syntactical reconstruction is shown by the dash-dotted curve in Fig. 1. The predicted Mach bands are relatively sharp and they occur at the four ramp-edge positions. In this example the amplitudes of the lines were determined by the actual cell responses. It is obvious that another weighting of these over all scales would lead to shallower or steeper edge slopes and other Mach-band shapes.

Figure 4 shows a more complicated signal together with the reconstruction. Most features of the input signal are also found in the reconstruction, but the reconstruction quality is not yet satisfying. A generally applicable analysis/reconstruction strategy is under study.

References

- [1] Bischof, W.F. and Caelli, T. (1988) Parsing scale-space and spatial stability analysis. *Comput. Vision, Graphics, and Image Proc.* **42**, 192-205.
- [2] du Buf, J.M.H. (1993) Responses of simple cells: events, interferences, and ambiguities. *Biol. Cybernet.* **68**, pp. 321-333.
- [3] du Buf, J.M.H. (1993) Ramp edges, Mach bands, and the functional significance of the simple cell assembly. *Biol. Cybernet.*, in print.
- [4] Venkatesh, S. and Owens, R. (1990) On the classification of image features. *Pattern Recogn. Lett.* **11**, 339-349.

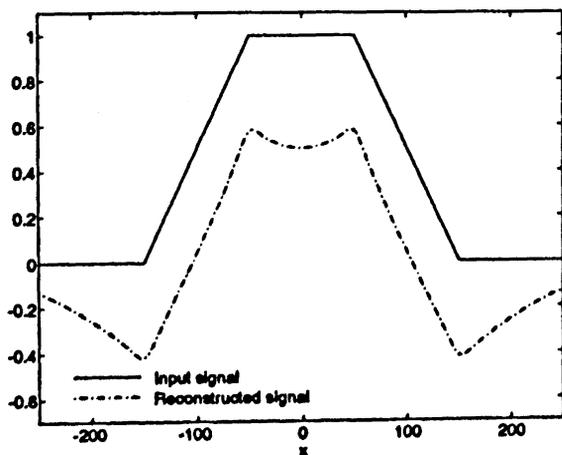


Figure 1: Trapezoidally shaped line and the syntactical reconstruction based on modulus maxima in the Gabor scale space.

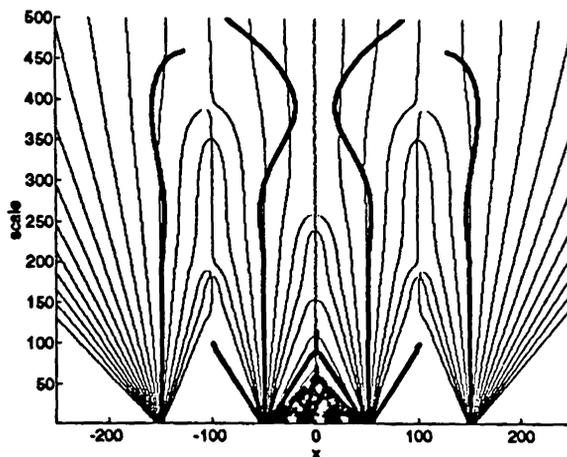


Figure 2: Gabor scale space with the maximum modulus (bold lines) and the zero crossings of the real and imaginary parts of the responses.

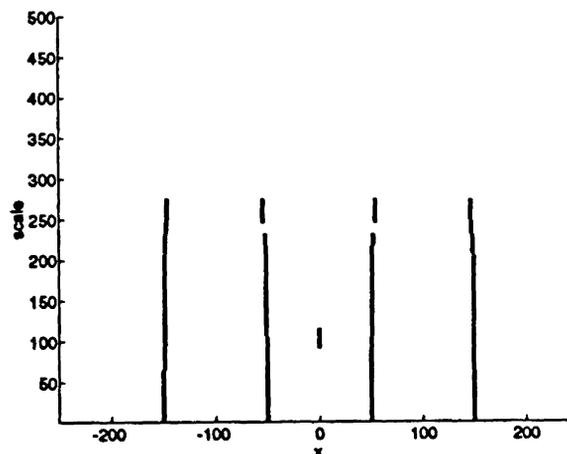


Figure 3: Stability of the maximum modulus position as a detection criterion. Only these moduli are used in the reconstruction (Fig. 1).

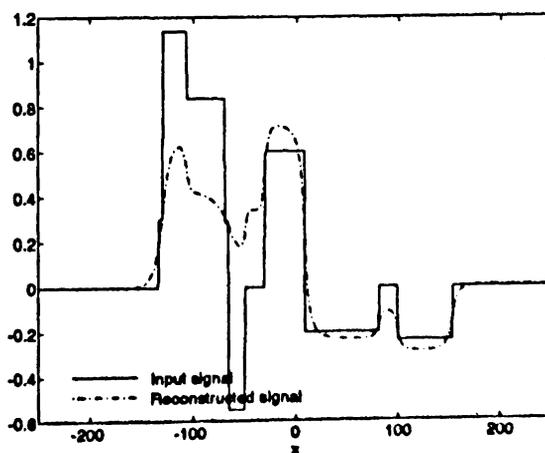


Figure 4: Another example of a reconstruction from the maximum modulus of the scale space representation.