Symbolic Pointillism: Computer Art motivated by Human Brain Structures

Norbert Krüger[1] and Florentin Wörgötter[2]

[1]Media Lab,
Aalborg University Copenhagen, Denmark

[2]Computational Neuroscience
Stirling University, Scotland

December 6, 2004

Abstract

We introduce a new kind of computer art that is motivated by cortical structures in the human visual system. It is related to the sub–group of the impressionist art movement called 'Pointillists'. However, while Pointillism visualises and makes use of processes that have been associated to the human eye, Symbolic Pointillism in addition makes cortical processes explicit. The visual representations underlying this art have been developed during a project that aims at the transfer of functional aspects of human vision to artificial systems. They have been applied in such an artificial vision system and in a sound/vision installation.

Introduction

Pointillism is a style of art that was part of the impressionist movement. This movement aimed to paint images not in a realistic way but in a way 'humans perceive the world'. The sub–group of impressionists called Pointillists (most notably Seurat and Pissarro) decided to base their paintings on small colored dots (or even small oriented patches). The observer then constructs the image by mixing these dots viewing them at a distance. Perceived reality is, thus, a concept constructed by the observer.

It is by now known that the human visual system operates by a similar principle. First, condensed and meaningful image information (like orientation, color, distance, form, shape) is generated by localized processors (neurons). These neurons are highly connected and communicate with each other. The need to reduce the cost of this communication process drives the condensation to meaningful information in the first place. As a consequence, it is almost as if those nerve cells would in the end communicate symbols. It seems to be this specific neuronal communication process by which our perceived 'reality' is generated.

Having the aim of building technical systems with similar power and structure to the human visual system, we invented an image representation that consists of abstract, symbolic primitives (described in the following section). We include, for example, information about color, orientation and contrast transition (see figure 1a). As found in the Nobel Prize winning work of Hubel and Wiesel [7], in the human visual system such information is coded in so

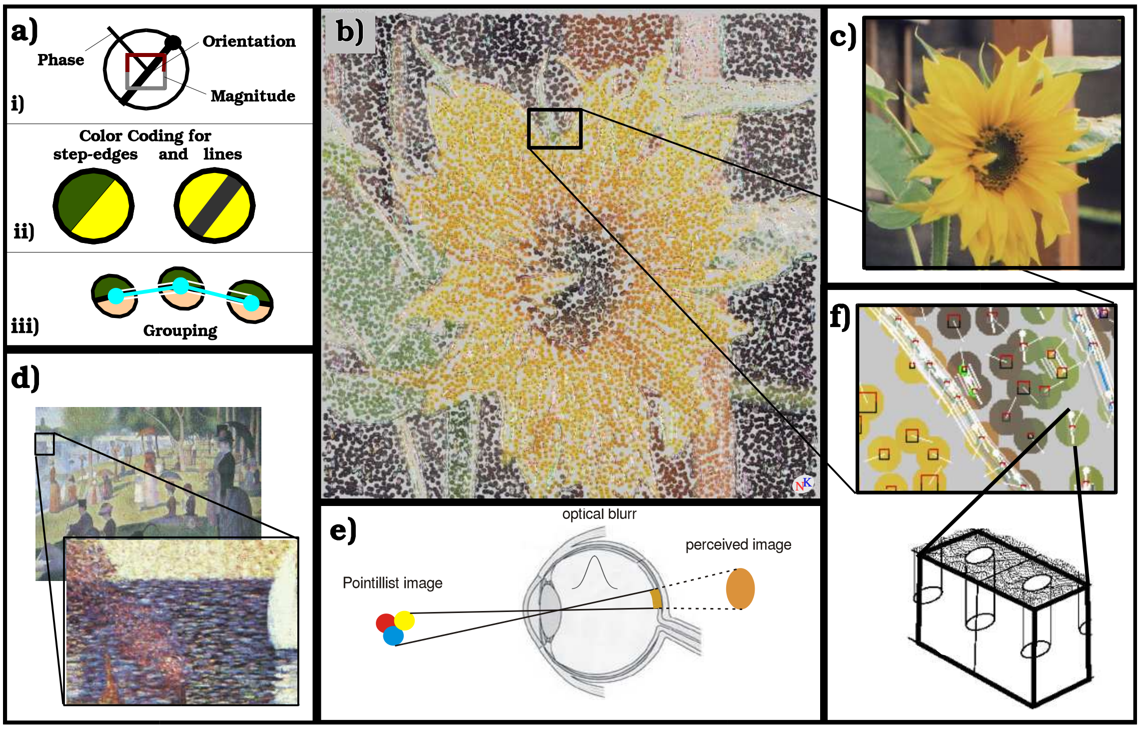

Figure 1: **a):** Schematic Icons the Symbolic Pointillistic image consists of. **b)** Symbolic Pointillistic image. **c)** Original image. **d)** Magnification of the Pointillist painting 'A Sunday Afternoon on the Island of La Grande Jatte' by Georges Seurat. **e)** The impression of color arises in Pointillist paintings and also in our images from the fact that the optical blurring of the eye leads to a merging of adjacent dots when viewing them from a distance. **f)** Magnification of the Symbolic Pointillistic image in b) where the last step depicts the schematic representation of a primitive. The functional aspects covered by the small icons inside are explained in c). This is meant to be similar to the information captured by a cortical hyper-column (see figure 4). Grouping processes are depicted by linking the small orientation markers to elongated lines.

called 'hyper-columns'. Moreover, communication processes are established across hyper-columns, for example by grouping of visual events (see, e.g., [18, 13]). Such a process has also been realized in our technical system where groups are represented by linking lines (see figure 1a,iii). In this way, our electronically generated paintings (that we have called 'symbolic pointillistic images' (see figure figure 1b and 2) represent aspects of image processing in the human brain in a schematic way.

The image representation is not designed for a specific application domain but makes use of generic principles of visual coding in human and computer vision. As a consequence, our image representation is widely applicable. We will discuss their role for the modeling of mid–level vision in artificial visual systems and for audio-visual mappings.

Symbolic Primitives and Early Visual Feature Extraction

In our artificial system we process a biologically motivated image representation that codes different aspects of visual information in a condensed way (for details, see [11]). Condensation has two aspects. Firstly, the

Figure 2: Two images of Symbolic Pointillism: 'Bunch of Margerets' and 'Baboon'

original image information is reduced in terms of bandwidth. Secondly, the rather unspecific pixel information is transformed to attributes of higher semantic.

Position and Orientation: To determine the position of the primitives we look for locations in the image where the magnitude of the response of a set of edge-detection filters [5] has local maxima. To avoid the occurrence of very close line–segments produced by the same image structure we also model a competition process between the primitives. Basically, for each primitive position it is checked whether another primitive exists with a position closer than a given threshold distance. If that is the case, the position with lower magnitude is dropped (for details see [10, 11]). Finding of suitable positions is a sophisticated task and is also part of a cruicial transformation process from a signal–based to a symbol–based representation. Once the positions of the primitives are determined, the orientation computed from the filter response at the found position is associated to the primitive. For edge–like structures, it is displayed by a local line segment (see figure 1a,i).

Contrast Transition: The contrast transition is displayed by a small line starting from the center of the primitive (see figure 1a,i). Using contrast transition, we can for example distinguish between lines and edges. Contrast transition is coded as the phase of a filter triplet (see [5]) at the primitive position. It refers to the kind of grey level structure existent at the local image patch (see also [6, 8]). This is displayed in figure 3a. For example a dark/bright edge is associated to a phase of $\pi/2$ while a bright line on a dark background has an associated phase of 0. However, as visible in figure 3a the phase codes a whole continuum of possible grey level structures by one parameter.

Color: Color is processed by averaging colour pixels in coincidence with their edge structure: For example, for edge like structures we distinguish between the colour on the left and right side of the edge (see figure 1a,ii).

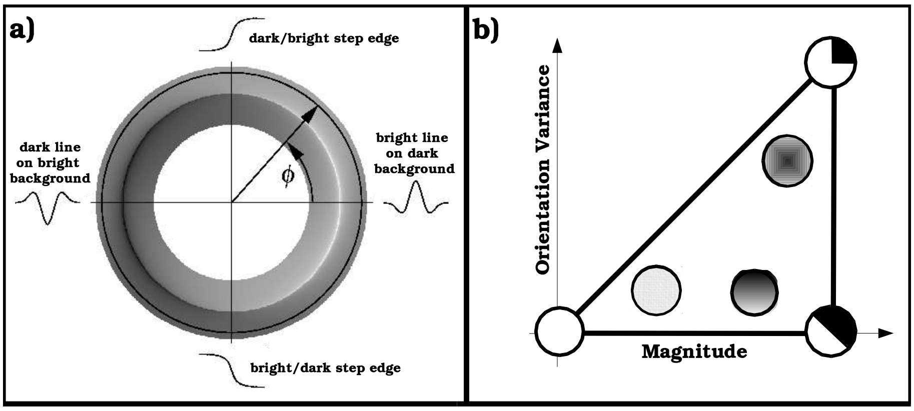

Figure 3: **a)** The continuum of phases (indicated by ϕ) taking values between $-\pi$ and π correspond to a continuum of oriented grey-level structures as expressed in the changing circular manifold (sub–figure a) is based on a figure in [4]). **b)** The likelihood of a local image patch to be a homogenous image patch, an edge or a junction can be visualised as a triangle with corners representing ideal patterns. Points inside the triangle represent structures that are only with a certain likelihood categorizable as ideal homogenuous image patches, edges, or junctions. For example, there is a slight texture on the patch close to the lower left corner which produces a filter response with low but measureable magnitude and orientation variance or the structure close to the upper corner has some resemblance to a junction. In this triangular representation distances from the corners represent the likelihood of the structures being of the ideal type. This is used for the formulation of confidences indicating such likelihoods in [9]. Note that figure 3b is thought to be a schematic description. The exact positioning of patches in the triangle depends on two parameters (for details see [9]).

Moreover, in case of a line structure (i.e., a primitive with phase close to 0 or π) we code also the colour of the middle strip (see figure 1a,i).

Magnitude and Orientation Variance: The 'homogeneousness', 'edge–ness' or 'junction–ness' of a local signal patch is an important descriptor that, e.g., allows for judging the applicability of concepts such as orientation and contrast transition. For example, a homogeneous image patch has no orientation or contrast transition (but has a color). Such a descriptor can be derived from two parameters computable from a local image patch: The local magnitude (which is a measure for dark/bright differences) and the variance of the local orientation. In [9] we showed that we can arrange different image structures in a triangle spanned by these two measures in which the corners correspond to ideal homogeneous image patches, edges, and junctions while points within the triangle represent image structures according to their distance to such ideal structures (see figure 3b). The orientation variance is displayed as the length of the diagonal of a square while the magnitude is displayed as the grey value of the lower part of the square (see figure 1a,i).

Quantification of Condensation: A Primitive contains a set of at most 15 parameters (described above) for the encoded information: 2 parameters for the position, 1 for orientation, 1 for contrast transition (phase), 9 parameters for colour (left, middle and right) and 2 parameters for the magnitude and the orientation variance. Note that only for line structures all 9 colour parameters are used while for the more frequent edge structures only 6 parameters are regarded. A primitive represents an image patch of appr. 12x12 pixels, each of which encodes

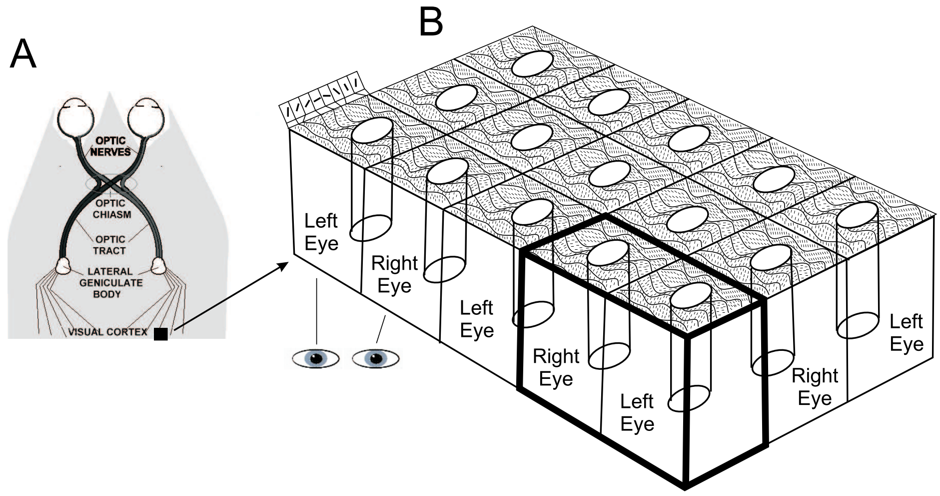

Figure 4: **a)** Primary visual pathway and schematic location of a hyper-column (black box), which corresponds in reality to about 1 mm^2 of cortical surface. **b)** Schematic diagram of a hyper-column (thick lines) embedded in the visual cortex. Each hyper-column represents a small location in visual space. Vertically to the surface neurons share similar response properties, whereas their responses differ when moving horizontally on the surface. Information from both eyes is represented in adjacent slabs of the cortex. Each slab contains neurons that encode different orientations (depicted by tiny lines on the surface) but also all other important visual features such as local motion and stereo. In the cylinder-shaped part mainly color is processed. Note, the actual cortical structure is less crystalline than suggested by this diagram.

three color values. Thus, each Primitive of 15 parameters represents a part of the image which contains a total of $3 \times 12 \times 12 = 432$ pixel values. Therefore, the primitives condense the image information by more than 96%.

Hyper-columns in the Visual Cortex

The above–mentioned visual modalities are processed at early stages of visual processing. Hubel and Wiesel [7] investigated the structure of the first stage of cortical processing that is located in an area called 'striate cortex' or V1 (see figure 4). The striate cortex is organized like a continuous, but distorted map of the visual field (retinotopic map). This map contains a specific repetitively occurring pattern of substructures called hyper-columns. Thus, a hyper-column represents a small location of visual space and the neurons in such a hyper-column represent all important aspects of this spatial location; ideally all orientations, all colors, the complete distance-information (disparity), etc. To be able to achieve this in an orderly manner, hyper-columns themselves are subdivided into "columns" and "blobs". The blobs contain color sensitive cells, while the columns represent the continuum of orientations (see figure 4b). Here one observes that the orientation columns are organized in an ordered way such that neurons representing similar orientations tend to be adjacent to each other. However, it is not only orientation that is processed in an orientation column but the cells are sensitive to additional attributes such as disparity, contrast transition and the direction of local motion (see [19]). Even specific responses to junction–like structures have been measured [16]. Therefore, it is believed that in the striate cortex basic local feature descriptions are processed similar to the feature attributes coded in our primitives.

However, it is not only local image processing that is going on in early visual processing. As mentioned above, there occurs an extensive communication within visual brain areas as well as across these areas. The communication process leads to the binding of groups of local entities (see, e.g., [18]). In [12] we described a self–emergence process in which groups organize themselves based on statistical regularities. In our images, primitives of the same group are represented by links of the same color (see figure 1a,iii).

Pointillism and Symbolic Pointillism

Why did we call our approach 'Symbolic Pointillism'? Historically pointillism (see figure 1d) has been interpreted as a process which happens "in the human eye".

> Seurat applied ... the Impressionist principle of using pure colors and allowing the mixture to be made in the spectator's eye.[15].

Indeed, it is the optical blurring of the eye that leads to the merging of colors when viewing these paintings from a greater distance (see figure 1e). As a consequence, although pointillists made use of only a small number of basic colours they can generate a huge variety of colour impressions. Visual blurring also occurs to a certain extent in symbolic pointillistic images. As in pointillistic images, people like to change the distance while observing the images. Taking a large distance, the symbolic icons merge into continuous forms while they become distinguishable when the observer approaches.

There exist two main differences between pointillsitic and symbolic pointillistic images. Firstly, symbolic pointillistic images are computer generated from photographs while pointillists were actually painters. Secondly, in symbolic pointillism we do not focus on retinal processes but rather on those that happen at the cortical level.

At the time of Seurat, theories were discussed by physiologists stating that the retina performs a process similar to additive color mixing [17]. By now it is clear that processing stages upstream of the retinal photo-receptors are more complex and cannot be explained only by color mixing (see, e.g., [14]). Basically, the whole visual machinery with its main components coded in cortical areas is involved. In addition to the blurring-induced effect mentioned above, our images capture the response properties of assemblies of cortical cells organised in the 'hyper–columns'. This is done in a symbolic way, depicted by icons for every primitive (see figure 1f). Since the goal behind our symbolic pointillist representation is to achieve advanced image processing properties, we needed to encode cortical information. Only at this level, integration, which is one of the first advanced processing steps, can be done in a meaningful way by, for example, grouping these Primitives into higher level entities (like elongated lines, see figure 1a,iii). Interestingly, in some few pointillistic images (for example Pissaro's painting 'Woman in Orchard') there is a step towards higher semantic entities by using, instead of coloured dots, small oriententTed coloured line segments as basic painting entity.

While observing pointillistic or symbolic pointillistic images, the whole visual machinery at all levels of processing is used. However, levels of cortical processing are made explicit to a wider extent in symbolic pointillistic images.

Modeling Mid-level Vision

We designed our Primitives with the aim to model intermediate stages of human vision. In the human brain a complex feature extraction process is realized in which different modalities are processed and communicate with each other. This feature extraction process forms the input to higher cortical areas where complex tasks such as object recognition, navigation or grasping are realized. It is known that, for example due to noise in the images, all low level features (orientation, color, edges, etc.) face the problem of an extremely high degree of vagueness and uncertainty [1, 13]. This makes a more complex feature extraction mechanism necessary and mainly *integration* across low-level features helps to disambiguate and correct the unreliable first bits of visual information. Such integration principles are realized in the human visual system allowing for visual representations necessary for actions with high precision and certainty. Similar integration processes are also possible and have been implemented in our artificial vision system (see, e.g., [11, 12]). This has become possible because symbolic primitives represent a *condensed* description of a local image patch, that contains the relevant information for integration. Condensation is important for integration because it limits data flow and data processing to a manageable level in a similar way as the number of long-ranging connections between cortical neurons is much smaller than those needed for short distance communication. The system has been developped to a large extent within the European project ECOVISION [3].

Audio-Visual Mappings trigger Voice Exploration

The Primitives have also been used in an audio-visual installation [2]. In this installation (see figure 5) the sound produced by a person is recorded and certain sound features are extracted. These sound features are then used to alter a symbolic pointillistic image. The altered visual feedback is then observed by the individual who can in turn influence his/her attempts to produce sound. In this way the perception–action loop (afferent/efferent loop) is closed which encourages people to explore their sound making capabilities. The condensed representation of the visual attributes in our primitives has proven to be essential since it allows for the mapping of extracted sound features to meaningful visual features. Psychophysical experiments were able to demonstrate that individuals become encouraged by such audio-visual loops to explore their voice and sound making capabilities. The installation has been presented at various occasions at Aalborg University, Denmark.

Figure 5: By making different kind of sounds the visual display becomes altered and feeds back to the person making the sound. This perception-action loop encourages people to explore their sound making capabilities.

Summary

We have introduced a novel form of computer art called ’symbolic pointillism’ that has been motivated by human vision. Symbolic icons represent visual information in a condensed way and make cortical stages of visual processing explicit. We pointed to the relation between ’pointillism’ and ’symbolic pointillsim’ and described two applications of our visual representations.

Acknowledgement: We would like to thank Helen MacGregor and Christina Griebner for taking the photos from which we have computed the symbolic pointillistic images ’Sunflower One’ and ’Bunch of Margarets’ (1b and figure 2,left). We thank M. Christensen, G. Bogason, M.F. Hansen, Karsten Broogaard, N. Johanen and Stefania Serafin for their work on the audio-visual mappings. We also would like to thank Kirstin Lyon for proofreading.

References

[1] Y. Aloimonos and D. Shulman. *Integration of Visual Modules — An extension of the Marr Paradigm.* Academic Press, London, 1989.

[2] Alastair Dalton. Painters inspire perfect technique to see yourself singing in tune. *The Scotsman, January 19*, page 8, 2004.

[3] ECOVISION. Artificial visual systems based on early-cognitive cortical processing (EU–Project). *http://www.pspc.dibe.unige.it/ecovision/project.html*, 2003.

[4] M. Felsberg. Optical flow estimation from monogenic phase. In B. Jähne, E. Barth, R. Mester, and H. Scharr, editors, *Complex Motion, Proceedings 1st Int. Workshop*, Günzburg, 12.-14.10. 2004.

[5] M. Felsberg and G. Sommer. The monogenic signal. *IEEE Transactions on Signal Processing*, 49(12):3136–3144, December 2001.

[6] G. H. Granlund and H. Knutsson. *Signal Processing for Computer Vision.* Kluwer Academic Publishers, Dordrecht, 1995.

[7] D.H. Hubel and T.N. Wiesel. Anatomical demonstration of columns in the monkey striate cortex. *Nature*, 221:747–750, 1969.

[8] P. Kovesi. Image features from phase congruency. *Videre: Journal of Computer Vision Research*, 1(3):1–26, 1999.

[9] N. Krüger and M. Felsberg. A continuous formulation of intrinsic dimension. *Proceedings of the British Machine Vision Conference*, pages 261–270, 2003.

[10] N. Krüger, M. Felsberg, and F. Wörgötter. Processing multi-modal primitives from image sequences. *Fourth International ICSC Symposium on ENGINEERING OF INTELLIGENT SYSTEMS*, 2004.

[11] N. Krüger, M. Lappe, and F. Wörgötter. Biologically motivated multi-modal processing of visual primitives. *The Interdisciplinary Journal of Artificial Intelligence and the Simulation of Behaviour*, 1(5):417–428, 2004.

[12] N. Krüger and F. Wörgötter. Multi modal estimation of collinearity and parallelism in natural image sequences. *Network: Computation in Neural Systems*, 13:553–576, 2002.

[13] N. Krüger and F. Wörgötter. Statistical and deterministic regularities: Utilisation of motion and grouping in biological and artificial visual systems. *Advances in Imaging and Electron Physics*, 131, 2004.

[14] M.W. Matlin and H.J. Foley. *Sensation and Perception (4th edition))*. Allyn and Bacon, 1997.

[15] Seurat. *Fifty Colour Plates.* Phaidon Press, 1965.

[16] I.A. Shevelev, N.A. Lazareva, A.S. Tikhomirov, and G.A. Sharev. Sensitivity to cross–like figures in the cat striate neurons. *Neuroscience*, 61:965–973, 1995.

[17] H. von Helmholtz, editor. *Handbuch der physiologischen Optic.* Hamburg & Leipzig: Voss, 1866.

[18] R.J. Watt and W.A. Phillips. The function of dynamic grouping in vision. *Trends in Cognitive Sciences*, 4(12):447–154, 2000.

[19] R.H. Wurtz and E.R. Kandel. Perception of motion, depth and form. In E.R. Kandell, J.H. Schwartz, and T.M. Messel, editors, *Principles of Neural Science (4th edition)*, pages 548–571. 2000.