Invited Paper
International Conference on Information Sciences, Signal Processing and its Applications
12–15 February 2007. Sharjah, United Arab Emirates (U.A.E.)

NOVEL EXPERT-BASED APPROACH TO IMAGE ANALYSIS

Charles M. Falco
University of Arizona, Tucson 85721 USA
falco@u.arizona.edu

ABSTRACT

The hands and minds of all painters are intimately involved in the creative process, making paintings intrinsically complex to analyze. Although tracing projected images is known to have become a common technique by the 19th century, earlier use of optics has been difficult to identify and analyze, hindered also by the lack of interaction between art historians and scientists. In spite of this difficulty, the painter David Hockney and I recently identified optical evidence within a number of paintings demonstrating artists as early as Jan van Eyck (c1425) used optical projections as aids for producing portions of their images. While making these discoveries, Hockney and I developed fundamentally new insights into image analysis. As discussed in this paper, I am now applying these new insights to problems in computerized image display and analysis.

1. INTRODUCTION

In reviewing the field of computerized image analysis, Rama Chellapa wrote[1]:

> *"The enhancement problem is made more difficult because image quality is decided by the not-so-well-understood human visual system…*
>
> *"The overall quality of image-restoration algorithms benefits greatly from: Better modeling of...the human observer and visual system…*
>
> *"Clearly, because images are viewed and evaluated by human observers, a realistic image-coding technique should incorporate characteristics that are peculiar to the human visual system.*
>
> *The main reason for lack of activity in this research area is perhaps the complexity of visual models and the interdisciplinary nature of the field…*
>
> *"The incorporation of complete human visual properties into the design of image-compression techniques should prove to be a challenging research area for the future."*

Although Chellapa made these observations fifteen years ago, this lack of understanding of how to exploit the human visual system for the various aspects of computerized image recognition remains a difficult obstacle today. The work reported here addresses this issue with a fundamentally new approach to computerized image analysis. This approach exploits the fact that the greatest artists have a remarkable understanding of how the features and context of images affect the recognition process in humans. Examples are given of initial work that demonstrates progress made to date in areas that are appropriate for informing the design of future automated image display and recognition systems.

2. BACKGROUND

2.1. Recent visual discoveries

Recently, the painter David Hockney reported visual discoveries within some of the best-known paintings of European art that affect long-held understandings of the development of Western art of the past 600 years[2]. In a collaboration combining the expertise and visual skills of one of the world's greatest artists[3] with the analytical skills of an optical physicist, we then developed the foundations of a new methodology for extracting information from complex, optics-based images[4,5,6,7,8].

Briefly, we showed that certain features within very well-known paintings (e.g. the chandelier in *The Arnolfini Marriage* by Jan van Eyck, as shown in Fig. 1) are based on optical projections. We determined that these optically-based elements of the paintings are "photorepresentations"[9]. Our discoveries show that optical projections were being used by artists over 150 years before Galileo brought an optical instrument, the telescope, to wide attention.

In the context of computerized image analysis, after an image is captured by a lens-based system, subsequent processing, including feature extraction, edge detection, image compression, etc., maintains the original encoding provided by the lens, as a flat field with an optically-imposed set of vanishing points. Images of interest can now contain over a million pixels, with a continuing drive for greater resolution. However, the encoding of these images is a fundamentally imperfect representation of human vision. Instead, when images are presented to observers in ways that mimic the way evolution has programmed our brains to function[10], humans can recognize images of remarkably low resolution. To

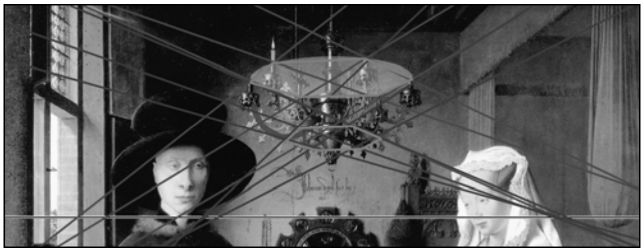

Figure 1. Jan van Eyck, *The Arnolfini Marriage*, 1434 (detail showing approximately 25% of the 81.8×59.7 cm painting). A summary of the evidence that the chandelier in this painting is based on optical projections is given in Ref [7].

demonstrate this, Fig. 2 is a painting by Vincent van Gogh.

The central feature of Fig. 2 is a person whose head occupies less than 5% of the surface area of the painting (subtending an angle of only ~1^o at a normal viewing distance). An enlarged detail of the painting reveals that van Gogh created the face using only three colors and with fewer than ten strokes of a rather broad brush (~5 mm). Yet, in spite of the "low resolution" of this image, the artist created a figure that viewers immediately recognize as a young woman standing in a field of flowers. The subject, Marguerite Gachet, was 19 at the time van Gogh made this painting[11,12].

One modern image analysis task involves observers quickly and accurately answering questions such as, "Is the feature on the computer monitor the same as one whose shape I studied in a 3D situation?" Reported here are initial results from a promising alternative approach to the current use of high resolution, photographically ideal images for such a task.

2.2. Paintings *vs.* photographs

As a consequence of the way they were constructed, optics-based paintings such as Fig. 1 are much more complex than photographs. However, in spite of this complexity, we have been able to extract quantitative evidence from such images even though they were produced by hand and are composites containing both optics-based and non-optics-based features. Our discoveries[2,3,4,5,6,7] have revealed that highly influential artists began using optical projections as aids for producing some of their paintings early in the 15th century, at the dawn of the Renaissance.

In addition to van Eyck (c1430), we have found optical evidence within works by later artists, including Bermejo (c1475), Holbein (c1530), Caravaggio (c1600), de la Tour (c1650), Chardin (c1750) and Ingres (c1825)[7], demonstrating a continuum in the use of optics by artists, along with an evolution in the sophistication of that use.

Figure 2. Vincent van Gogh, *Mlle. Gachet in Her Garden at Auvers-sur-Oise*, 1890 (46×55.5 cm).

However, even for paintings within which unambiguous, quantitative evidence exists of the direct use of optical projections for producing certain of the features, it does not mean that these paintings are effectively photographs. Because the hand and mind of the artist are intimately involved in the creative process, understanding these images requires more than can be obtained from only applying the equations of geometrical optics. For example, the full painting of Fig. 1 contains a small dog at the feet of the subjects. Since it is very unlikely any such dog would have stood sufficiently motionless long enough for van Eyck to capture details from its projected image, we can be reasonably certain that elements of this painting, such as the dog, were made without use of direct optical projection.

Although I only briefly address it in this paper, no less important for understanding the evolution of post-c1425 painting, as well as certain modern applications of image analysis, is the indirect use of optics. Unlike an image projected onto film, the human eye constantly adjusts its aim and focus as the mind constructs the scene it is viewing. As a consequence, humans do not simultaneously see part of a scene in focus and part out of focus. Hence, a simple example of the indirect use of optics is if an artist has painted a distant portion of a scene out of focus, replicating the depth-of-field of an image projected by a lens. Although modern humans have seen this effect countless times in the form of photographs, in movies, and on television, it is not an effect that is part of natural human vision.

The fact that psychology is as intimately involved in vision as is the simple geometrical optics of the eye occupies a significant part of Ibn al-Haytham's seven-volume treatise on optics[13,14], the first time this topic was first addressed in a modern scientific fashion. Al-

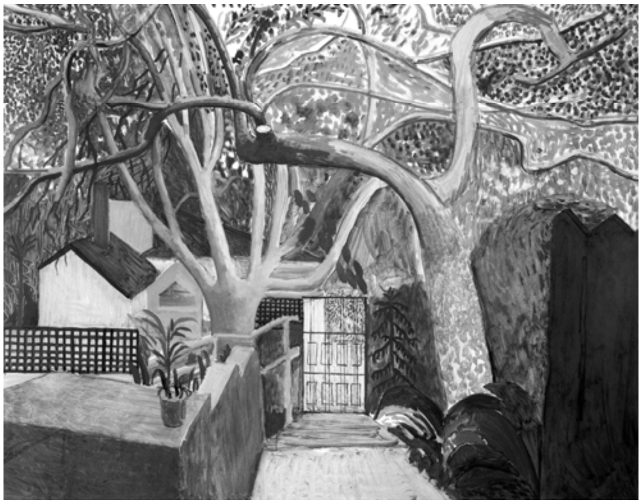

Figure 3. Painting of the scene as viewed from Hockney's studio door. David Hockney, ©2000.

Haytham's landmark work *Kitāb al-Manāzir [Book of Optics]* was translated into Latin in the early thirteenth century[15], and had a profound influence on European intellectuals, including figures as diverse as the writer Geoffrey Chaucer, the theologian John Wyclif[16], and the scientific work on optics of Bacon, Pecham, and Witelo[17]. Al-Haytham's work was republished in Latin in 1572, after the advent of the printing press, and is explicitly referenced in the writings on optics by Kepler, Snell, and Fermat[14].

Figure 3 is a painting that David Hockney made of the scene outside his studio door subsequent to the start of research that discovered the use of optics by the Old Masters[2]. The way he composed this painting provides important insights into how the visual skills of artists can be exploited for informing computerized imaging tasks, as discussed below.

In contrast to Hockney's painting, the image in Fig. 4 shows that it requires a composite of six photographs (and hence six sets of vanishing points) taken with a moderately wide angle lens[18] to capture the various important features contained within that painting. But, even though this composite photograph contains far more "data" than does the painting, and even though we are conditioned to think of photographs as accurately representing reality, they are not necessarily optimal for conveying visual information in a form most readily recognizable to humans.

As just one example, note how the central feature of the painting, the gate at the end of the sidewalk, is much more prominent in Hockney's representation of the scene than it is in the composite photograph. His exaggerated scale and placement of that gate is how my experience of the scene tells me most observers would remember it if they were

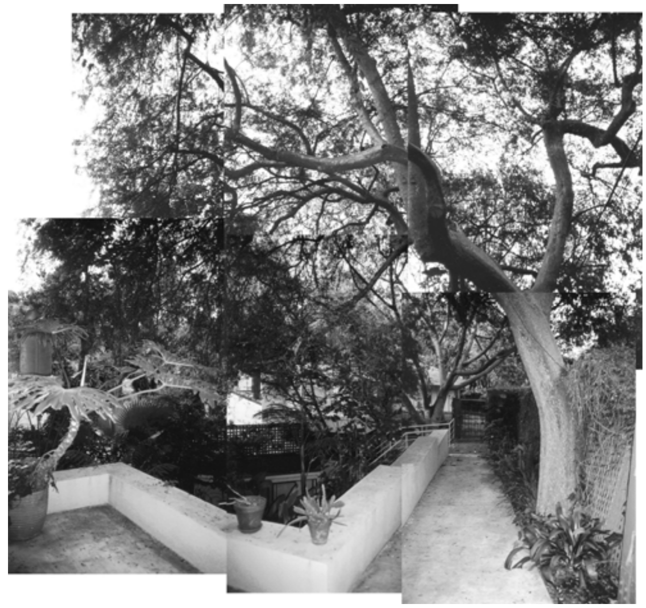

Figure 4. Composite photograph of the same scene represented in Fig. 3, made up of six overlapping semi-wide angle photographs[18].

asked about it some time later. Restating this, my premise is that if people who had visited Hockney's studio were later shown either the composite photograph or the painting, more of them would identify the original location of the scene more quickly and accurately from the painting, even though the painting is "distorted" and consists of far fewer "data" than in the high resolution photograph. A very rough estimate, based on experimenting with jpegs at various levels of compression, is that Hockney's painting of the scene contains no more than 5% the data of the composite photograph.

2.3. Beyond simple lens-based imaging

Incorporating appropriate insights from the highly-developed visual skills of renowned artists like David Hockney never has been previously done in the field of computerized image analysis, even though understanding and exploiting the human visual system has been recognized for quite some time by the image analysis community as important for making progress[1]. This paper describes some initial results toward that end.

The van Gogh and Hockney painting examples illustrate that humans acquire and process visual information in a more complex way than does a simple camera lens. In spite of this, for over 170 years the direct image projected by a lens has remained the basis for approaches to image analysis. The examples given next show how two other imaging experts (i.e. highly skilled artists) used their keen understanding of human visual response to convey visual information to others. Insights

Figure 5. Photograph of the Houses of Parliament taken from a specific location described in the text.

Figure 6. Claude Monet, *Houses of Parliament (soleil couchant)*, 1905 (82×91 cm). Outline of skyline and river taken from Fig. 5.

derived from these and other examples have the potential to enable revolutionary improvements in a wide range of computationally-complex image processing applications.

3. QUANTITATIVE INFORMATION FROM QUALITATIVE IMAGES

3.1. Background

Figure 5 is a photograph of the Houses of Parliament in London, taken from a specific location on the bank of the river Thames 440 meters from the right edge of the tallest tower in the center of the photograph[19]. Traced over the photograph is the outline of Parliament and of the river edge.

Since photographs are two-dimensional representations of three-dimensional scenes, some information from the original scene inevitably will be missing. For example, in the case of Fig. 5 it is not possible to tell that the short tower immediately to our right of the tallest tower is significantly closer to the camera. Quantitatively, the left edge of the shorter tower is located 110 m closer to the camera than is the right edge of the tallest tower. The significance of this is that parallax will affect the amount of apparent separation between these two towers. The apparent separation will become larger if the camera is moved along the river to the left, and smaller if it is moved to the right. Because the distances to the two towers are significantly different, the effects of parallax will be relatively large. Hence, this visual separation would allow someone to later locate to fairly high precision the position along the river from which the photograph of Fig. 5 was taken. As will be shown next, in addition to other information, this fact has allowed me to positively identify for the first time the specific location where the French Impressionist painter Claude Monet stood when he made a series of nine non-lens-based Impressionist paintings of this scene over a two year period during 1903–05.

3.2. Claude Monet's *Houses of Parliament*

Figure 6 is one of Monet's paintings from his nine-painting series, over which is placed the outline traced from the photograph of Fig. 5. Because of the parallax between the two towers discussed above, we can identify that the location where Monet stood when making this painting is the same as the location of the camera for the photograph in Fig. 5. As can be seen, the overall fit of the outline to the painting is remarkable, with the exception of the tallest tower. The accuracy with which Monet was able to reproduce the complex features of the skyline shows that, in addition to his artistic acuity, he had excellent visual and technical skills. However, he clearly painted the tower both taller and narrower than it appears in the photograph.

The nine paintings in this series are all rather large (~81×92 cm). From the measured distance to Parliament from this location (362 m to the center of the line running between the two tallest towers visible in Fig. 5), it can be calculated that by standing 1.4 m back from the painting on its easel, the painting and the actual scene would have occupied identical visual angles. Hence, the significant discrepancy in the height and width of the tallest tower would have been readily apparent to Monet at the time he was making the painting.

To test this further, Fig. 7 shows the same fit to all nine of Monet's paintings in this series, produced in a variety of lighting conditions, and with slightly different compositions, over a period of two years. As can be seen, in all cases the overall fit is excellent, with the exception of the tallest tower, which Monet made taller and narrower

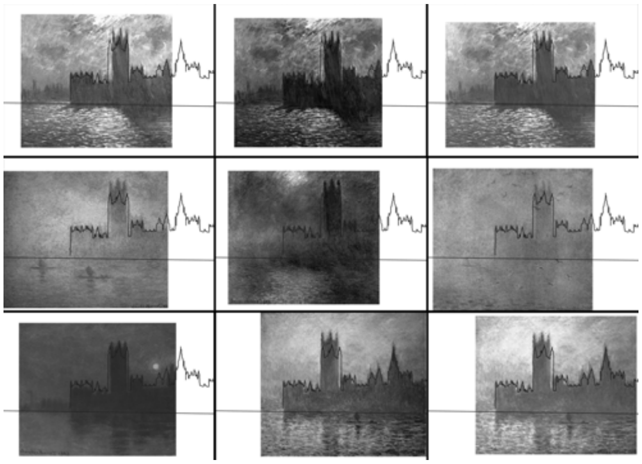

Figure 7. Claude Monet, *Houses of Parliament*, 1903–05 (composite; each of the nine paintings is approximately 82×91 cm).

than the actual tower appears when viewed from that location. Quantitatively, Monet made the tower on average 15.5% too tall (min. 10.2%; max. 23.6%) and 18.1% too narrow (min. 15.6%; max. 21.9%).

From this analysis, my conclusion is that Monet deliberately exaggerated the relative dimensions of the tallest tower. Rather than produce two-dimensional representations that were "photographically accurate," he altered the geometry in a specific way that his visual skill and artistic brilliance told him would best represent his view of the original scene to people who subsequently saw these paintings.

3.3. André Derain's *Big Ben*

Figure 8 shows a 1906 painting by another French Impressionist artist that my optical analysis shows was made from the same location as Monet used, but with the artist facing directly across the Thames (i.e. roughly 45^{o} to the right of the direction Monet faced). As in the previous example, the skyline overlaid on this painting is from a photograph taken from the same location as for Fig. 5[20]. Although the overall fit of the skyline to this painting is not as good as that to Monet's paintings, it is still apparent that the artist has represented the central feature in this painting, Big Ben, significantly taller than it should be for it to be "photographically accurate."

4. SIGNIFICANCE FOR COMPUTERIZED IMAGING

Hockney, Monet, and Derain are all renowned artists of the highest rank, with the technical skills to produce paintings that closely reproduce the optical perspective of a camera lens if that is what they desired to do. Since I was able to identify the precise locations each of these artists stood

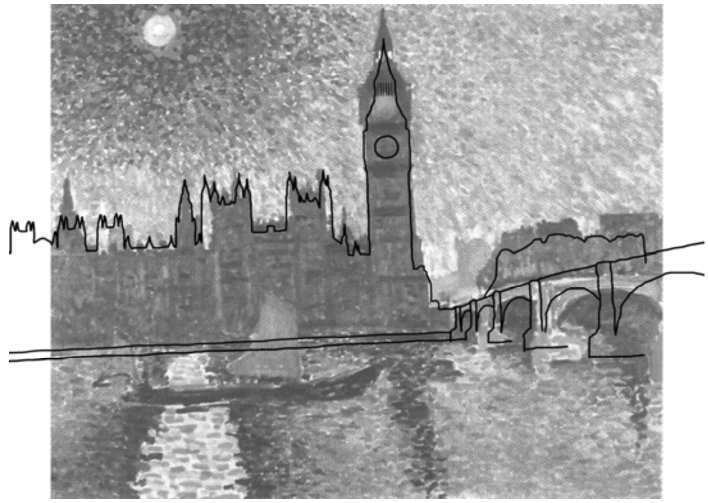

Figure 8. André Derain, *Big Ben*, 1906 (78×98 cm). The outlines of the skyline, bridge, and river were taken from a photograph similar to Fig. 5.

when making the paintings shown in this paper, the specific deviations from "photographic accuracy" in these examples provide us with key insights into what their visual skills informed them were the ways to represent these two-dimensional images to viewers. In the context of computerized image analysis, one future direction will be to apply this information to develop algorithms that automatically "distort" photographic images in specific ways that enable viewers to more quickly and accurately recognize visual information than they can from the as-captured images.

Although certain artists have visual skills of the highest possible level, art intrinsically is non-analytical. In spite of this seemingly-insurmountable difficulty for quantitative work, this paper shows initial results from a new approach to image analysis that successfully incorporates for the first time appropriate information from such artists with the tools of optical physics.

ACKNOWLEDGMENTS

I gratefully acknowledge David Hockney for the many invaluable insights I gained from him in our collaboration that investigated paintings from over 1000 years of European art. These insights provided the foundation for the research reported here, which is supported by ARO grant W911NF0610359 and DARPA grant NBCH1050008

REFERENCES

[1] Chellappa, Rama. *Digital Image Processing*. (IEEE Computer Society Press, 1992).

[2] David Hockney, *Secret Knowledge: Rediscovering the Lost Techniques of the Old Masters* (Viking Studio, 2001).

[3] Hockney, David (1937–)

"British painter, draughtsman, printmaker, photographer, and designer... by far the best-known and most critically acclaimed British artist of his generation."

The Oxford Dictionary of Art and Artists, Ian Childers, ed. (Oxford University Press, 1996).

"...His drawings and etchings are among the deftest of this century; posterity may well acclaim him the greatest of modern portraitists."

The Yale Dictionary of Art and Artists, Erika Langmuir and Norbert Lynton, eds. (Yale University Press, 2000).

[4] David Hockney and Charles M. Falco, "Optical Insights into Renaissance Art," Optics & Photonics News **11**, 52 (2000).

[5] Hockney, David and Charles M. Falco, "The Art of the Science of Renaissance Painting," *Proceedings of the Symposium on 'Effective Presentation and Interpretation in Museum'* (National Gallery of Ireland, 2004).

[6] Hockney, David and Charles M. Falco, "Optical Instruments and Imaging: The Use of Optics by 15th Century Master Painters," (Proceedings of Photonics Asia, SPIE, 5638, 1, 2005).

[7] Hockney, David and Charles M. Falco, "Quantitative Analysis of Qualitative Images," (Proceeding of IS&T-SPIE Electronic Imaging, SPIE, 5666, 326, 2005).

[8] Falco, Charles M., "Analysis of Qualitative Images," *Proceedings of the Irish Machine Vision and Image Processing Conference.* A. Amira, A. Bouridane, and F. Kurugollu, eds. (Queen's University Belfast, 2005).

[9] I use the word "photorepresentation" to avoid the reader drawing the incorrect conclusion that elements within these paintings are effectively photographs. Even through a projected image is on the surface in front of him, the artist, unlike a piece of photographic film, is free to use artistic judgment to trace portions of it exactly as projected, alter other portions to suit his taste, and ignore yet other parts of that projected image entirely. Consequently, these paintings are not simply composites of accurate tracings.

[10] The visual systems of primates have evolved over 30–50 million years to the point where the human visual cortex occupies over one-third of our brain mass. As a result, even 20,000-years ago humans were able to produce remarkably realistic images (e.g. the cave paintings at Lascaux, France). In contrast, cuneiform tablets from 5000 years ago show that even by that date our mathematical abilities had yet to advance beyond counting.

[11] Although not a controlled scientific study, I have tested a color version of this image on over twenty people. All quickly identified the central figure as female, and none gave an age for her of less than 18 or over 30.

[12] Saltzman, Cynthia. *Portrait of Dr. Gachet: The Story of a Van Gogh Masterpiece* (Viking Press, New York, 1998).

[13] al-Haytham, Abū 'Alī al-Hasan ibn al-Hasan ibn (Latinized as Alhazen or Alhacen). *Kitāb al-Manāzir [Book of Optics]* (Cairo, c1028–38).

[14] Lindberg, David C. *Theories of Vision from Al-Kindi to Keppler* (University of Chicago Press, 1976).

[15] Smith, A. Mark. *Alhacen's Theory of Visual Perception* (American Philosophical Society, 2001). Smith provides an insightful analysis of the 22 existing Latin manuscripts of the first three books of al-Haytham's text. It was in Latin, with translation interpretations, omissions and errors, that the text was studied by Europeans.

[16] al-Haytham and his optics are described in the c1275 epic poem *Romance of the Rose*, which for the next 300 years was one of the most widely-read works in the French language, and in Chaucer's c1400 *Canterbury Tales*, the first major piece of literature in the vernacular English language. The 14thC theologian John Wyclif, whose writings were an important precursor to the Protestant Reformation, used al-Haytham's writings on physical vision to help explain his own doctrine of spiritual vision.

[17] Sabra, A. I., "Ibn al-Haytham," in, *Dictionary of Scientific Biography*. Charles C. Gillispie, ed. (Charles Scribner's Sons, 1972).

[18] The photographs were made on 35 mm-format film (24×36 mm frame) with a 35 mm focal length lens.

[19] I obtained distances using Google Earth™, whose pointer provides a resolution of 0.01 arc-sec., corresponding to a distance resolution of better than 0.2 m.

[20] The buildings across the road to the right of Big Ben today are different than the ones that were there in 1906. Because of this, for this small portion of the skyline I substituted a tracing from a period photograph that had been taken from roughly, but not precisely, the same location. Hence, the portion of the skyline directly above the bridge is only approximate.