Bringing Pictorial Space to Life:
Computer Techniques for the Analysis of Paintings

Antonio Criminisi[1], Martin Kemp[2] and Andrew Zisserman[3]

[1] Microsoft Research Ltd., 7 J J Thomson, Cambridge, UK

[2] Dept. History of Art, University of Oxford, Oxford, UK

[3] Dept. Engineering Science, University of Oxford, Oxford, UK

http://research.microsoft.com/∼antcrim/

Abstract

This paper explores the use of computer graphics and computer vision techniques in the history of art. The focus is on analysing the geometry of perspective paintings to learn about the perspectival skills of artists and explore the evolution of linear perspective in history.

Algorithms for a systematic analysis of the two- and three-dimensional geometry of paintings are drawn from the work on "single-view reconstruction" and applied to interpreting works of art from the Italian Renaissance and later periods.

Since a perspectival painting is not a photograph of an actual subject but an artificial construction subject to imaginative manipulation and inadvertent inaccuracies, the internal consistency of its geometry must be assessed before carrying out any geometric analysis. Some simple techniques to analyse the consistency and perspectival accuracy of the geometry of a painting are discussed.

Moreover, this work presents new algorithms for generating new views of a painted scene or portions of it, analysing shapes and proportions of objects, filling in occluded areas, performing a complete three-dimensional reconstruction of a painting and a rigorous analysis of possible reconstruction ambiguities.

The validity of the techniques described here is demonstrated on a number of historical paintings and frescoes. Whenever possible, the computer-generated results are compared to those obtained by art historians through careful manual analysis.

This research represents a further attempt to build a constructive dialogue between two very different disciplines: computer science and history of art. Despite their fundamental differences, science and art can learn and be enriched by each other's procedures.

A longer and more detailed version of this paper may be found in [5].

1 Introduction

In the twentieth century, art and science were generally perceived as very diverse disciplines, with very few points of contact between them. Although there are signs that the schism is less sharp than it was, we are a long way from the situation that prevailed in the Italian Renaissance, when the distinction between those who practiced what we call art and science was not sharply drawn. At that time innovative artists were seen as ingenious people (credited with "ingenium"), capable of inventing their own systematic techniques for rational representation, designing new instruments and graphic tools, and striving to achieve their goals according to their own kind of "science". Leading figures worked not only as artists but also as engineers and scientists. Striking examples of such "Renaissance men" include Brunelleschi, Masaccio, Piero della Francesca, Albrecht Dürer and Leonardo.

The work in this paper – building upon that by Kemp [17] – aspires to show how, even today, scientific analysis and the study of art can interact and be mutually beneficial in achieving their goals. Novel and powerful computer techniques can help art historians to answer such much- debated questions as: is the geometry of Masaccio's *Trinity* correct? how deep is the *Trinity*'s chapel? what is the shape of the architectonic structure in Piero della Francesca's *Flagellation*? What is the shape of the dome in Raphael's *School of Athens*? The focus of this paper is to show

how computer graphics and computer vision can help give new kinds of answers to these as well as other interesting questions about the spatial structure of paintings.

All the paintings and frescoes that will be taken into consideration in this work share a great sense of perspective. Indeed, a perspectival structure of some elaboration is necessary if the proposed method is to yield meaningful results. Linear perspective was invented in the second decade of the fifteenth century in Florence by Filippo Brunelleschi. During the next decade it began to be used by innovative painters as the best technique to convey the illusion of a three-dimensional scene on a flat surface such as a panel or a wall. Masaccio, Donatello, Piero della Francesca, Domenico Veneziano and Paolo Uccello were amongst the first to experiment with this very new technique. In the seventeenth and eighteenth centuries a number of mathematicians such as Desargues, Pascal, Taylor and Monge became increasingly interested in linear perspective, thus laying the foundations of modern *projective geometry* [12, 21, 22]. Projective geometry can be regarded as a powerful tool for modeling the rules of linear perspective in a metrical or algebraic framework.

Over the past ten years projective geometry has become the basis of many of the most powerful computer vision algorithms for three-dimensional reconstruction from multiple views [11, 13, 9]. In particular, the work on single-view metrology [3, 15, 19] provides tools and techniques to compute geometrically accurate three-dimensional models from *single* perspective images. Those single-view techniques are applicable to all perspectival images (such as photographs) and are extensively applied, in this paper, to paintings which aspire to create a systematic illusion of space behind the picture plane, i.e. those which adhere to the canonical rules of linear perspective [1].

This paper, rather than just showing three-dimensional reconstructions designed on the basis of data extracted from paintings, presents novel and general methods which may be applied directly to any perspective image for a thorough analysis of its geometry.

We should constantly bear in mind that a painting is a creation that relies upon the artist's and spectator's imagination to construct a new or artificial world. It originates from the hands of an artist skilled in achieving effects in which the manipulation of orthodox perspective may be advantageous and in which accuracy may not be a paramount consideration. Therefore, any kind of purely geometric analysis must be carried out in a diplomatic and sensitive manner. In particular, before any geometric reconstruction can be applied it is necessary to ascertain the level of geometric accuracy within the painting and, by implication, the desire of its maker for perspectival precision.

In this paper we present some simple techniques for assessing the consistency of the painted geometry. This is done, for example, by using powerful techniques to analyse the location of vanishing points and vanishing lines, by checking the symmetry of arches and other curved structures and by analysing the rate of diminution of receding patterns (e.g. tiled floors and arched vaults).

If a painting conforms to the rules of linear perspective then it behaves, geometrically, as a perspective image and it can be treated as analogous to a straightforward photograph of an actual subject. Vision algorithms can then be applied to: generate new views of the painted scenes; analyse shapes and proportions of objects in the scene; complete partially occluded objects; reconstruct missing regions of patterns; perform a complete 3D reconstruction and an analysis of the possible ambiguities in reconstruction.

1.1 Alternative Methods in History of Art

The method we are advocating here needs to be set in the context of previous techniques for the analysis of the perspective of paintings. Together with the present method, there are now three main alternatives, which exhibit different sets of advantages and disadvantages.

Hand-made. The longest standing method has been to draw lines on the surface of paintings, or, rather, on the surfaces of reproductions of paintings (for obvious reasons). The linear analyses can be conducted either directly on a reproduction, or by using transparent overlays, and the results can be shown either superimposed on a reproduction or as separate diagrams. This latter form of presentation has been adopted by Kemp [17] amongst others. The analysis of a painting should preferably be performed on as large a reproduction as is available, ideally life-size (though this

[1] The multiple-view techniques developed in the Computer Vision and Computer Graphics communities such as [9] are not applicable to accurately reconstructing and analysing paintings because of the lack of stereopsis and scene measurements.

is rarely possible). The advantages of the hand-drawn analyses are: (i) the technique is congruent with that used by most artists themselves, who typically constructed their illusions of space through preparatory work involving straight edge and linear measures (sometimes also using dividers and compasses); (ii) in the initial stage of analysis, the hand drawn lines can work flexibly for experimental exploration of our intuitions about the structure of the depicted space.

The disadvantages are: (i) it is all to common to find thick lines drawn on small reproductions of big paintings, with resultant imprecision; (ii) it is easy to make errors, as when drawing a line through the point of intersection of two other lines, in which the extension of the resulting line will deviate progressively in response to the error in exact placement at the vertex of the intersecting lines. (iii) constructing diagrams for complex illusions is long-winded.

Traditional Computer Aided Design. More recently, data obtained from the analysis of paintings has been used to obtain a computer aided design (CAD) reconstruction of the depicted space, using standard programmes [8]. The advantages of the CAD Reconstructions are: (i) they depict spatial features with precision according to the classic rules of linear perspective; (ii) using techniques for rendering, they produce pictorial effects of light, shade, colour and (to a degree) texture akin to those in the original image; (iii) they can be used to produce animated fly-throughs and externalised views of the reconstructed spaces that are vivid aids to understanding.

The disadvantages are: (i) they require data to be extracted in advance from the painting, often in an artificially "tidied-up" manner, in order to work with the programme; (ii) they acquire a separate existence from the original images and may assume an aura of precision and conviction that is attractive but spurious.

The present method. The method advocated here still belongs to the wide spectrum of CAD applications; but, unlike traditional techniques, works directly from the surface of the painting, and does not, for the purpose of analysis, add any arbitrary data not embedded in the image itself. The advantages are: (i) alternative starting assumptions about the space in the image can be explored with ease and compared. All the alternatives can be parametrized in a rigorous, mathematical fashion; (ii) the internal consistency and inconsistencies of the spatial representation are laid bare, using information available directly from the image itself and involving no re-drawing; (iii) the textures and colours of the original image are retained; (iv) re-projection can scrutinise errors and allow for their correction; (v) degrees of inaccuracy can be estimated systematically across the surface of images; (vi) fly-throughs and externalised views of the space can be produced with ease; (vii) regions occluded by objects closer to the viewer can be systematically reconstructed.

The disadvantages are: (i) the power of the analysis may be excessive for the quality of the information which the artist entered into the original painting; (ii) it is applicable with profit only to those paintings which were constructed with sustained attention to perspectival rules; (iii) the quality of resolution is dependent upon the quality of the source image.

Some of the disadvantages of the present method may apply also to the previous two methods.

We like to think of the technique described in this paper as a more flexible Computer-Assisted analytical tool which, unlike other CAD applications, allows the user to analyse the geometrical information contained in the painting in a more rigorous way by exploring all the possible reconstruction alternatives systematically, avoiding implicit assumptions typical of the use of traditional geometric templates.

We maintain that the third method is generally superior for the analysis of complex perspectival images, and point to the first two advantages as empirically decisive in relation to the other two methods.

Correlating the Results with Historical Knowledge. The results obtained by these or other possible methods all need to be correlated by the historian with three other main bodies of evidence: (i) the *archaeology* of the painting; that is to say the physical evidence embedded in the work itself which reveal the constructional methods employed. These include incised lines (such as are apparent in Masaccio's *Trinity*), underdrawings detectable with such techniques as infra-red reflectography, and any *pentimenti* (changes of mind) visible in the surface of the painting; (ii) evidence from drawings by the artist and comparable artists about the methods they employed to construct perspectival spaces; (iii) the techniques for spatial construction available at the time the paintings were made, as recorded in published and unpublished treatises and diagrams.

The remaining sections of the paper are organized as follows: Section 2 describes some techniques to assess the accuracy of a painting's linear perspective. Section 3 presents algorithms to analyse patterns and shapes of objects in paintings. Finally, section 4 presents complete three-dimensional reconstructions of paintings, analyses the dependency of the reconstructed geometry upon the assumptions made and explores the possible reconstruction ambiguities. Throughout the paper the labourious analytical work of patient art historians is shown side by side with the results originated by applying our efficient and rigorous vision techniques to paintings.

2 Assessing the accuracy of a painting's linear perspective

As stated in the introduction, injudicious application of reconstruction techniques to paintings may lead to disastrous results. Our first task is to assess how well a painting adheres to the rules of linear perspective; in other words, how accurately its geometry represents that of a three-dimensional scene. This section presents some simple techniques used to perform this task.

Consistency of vanishing points. Under perspective projection (e.g. taking a photograph with a camera) lines parallel to each other in a real scene (e.g. the edges of a table, the tiles of a floor, or the edges of windows on a building facade) are imaged as converging lines on the image plane. The intersection point is called *vanishing point*. This holds for *any* set of lines as long as they are parallel to each other in the scene.

Therefore, one first technique for assessing the accuracy of a painting's geometry is that of analysing how well parallel lines converge into their vanishing point.

The rate of receding regular patterns. Assessing the consistency of the location of vanishing points or vanishing lines may not be sufficient to decide whether a painting obeys the rules of linear perspective. For example, in a drawing or painting of a regular pattern such as a tiled floor even though the edges of the tiles may consistently intersect in a single vanishing point, the rate of diminution of the tiles areas may be incorrect.

A detailed description of these two first techniques is omitted here (for lack of space) and may be found in [5].

A more interesting way of assessing the correctness of the geometry of a painting is that of comparing heights of objects, as described in the following section.

2.1 Comparing heights

Even in perspectivally constructed images the heights of figures might be varied by the artist according to the status of those represented. The *Virgin and Child* for instance, was sometimes accorded a larger scale relative to persons of lesser status, in a way that is not immediately apparent to the unaided eye. The heights of donors when they are depicted in paintings such as altarpieces was particularly subject to variation in relation to the holy figures in the same or adjacent spaces. Therefore, comparing the heights of people in a painting can prove interesting in order to ascertain their consistency with the perspective rules and to attempt to establish whether any disproportion is an intentional response to hierarchies of status.

The schematic in fig. 1 is used to explain how heights of people can be computed *directly* from perspective images. In fig. 1a we wish to compute the height of the man with respect to the height of the column. The column (or any other object in the painting) is used as a reference. The vanishing line of the ground plane (i.e. the horizon) has been computed (e.g. by joining two or more vanishing points corresponding to horizontal directions in the scene) and shown in blue. The line joining the base of the column (point $\mathbf{r}_b$) with the base of the man (point $\mathbf{x}_b$) intersects the vanishing line in the point $\mathbf{v}$. The line joining the top of the column (point $\mathbf{r}_t$) with the point $\mathbf{v}$ intersects the vertical through the man (dashed green line) in the point $\mathbf{i}$.

This construction has projected the chosen reference height onto the vertical through the man. In fact, the two lines $< \mathbf{r}_b, \mathbf{v} >$ and $< \mathbf{r}_t, \mathbf{v} >$ are images of parallel lines in space and the points $\mathbf{i}$ and $\mathbf{r}_t$ are at the same height from the ground, in space.

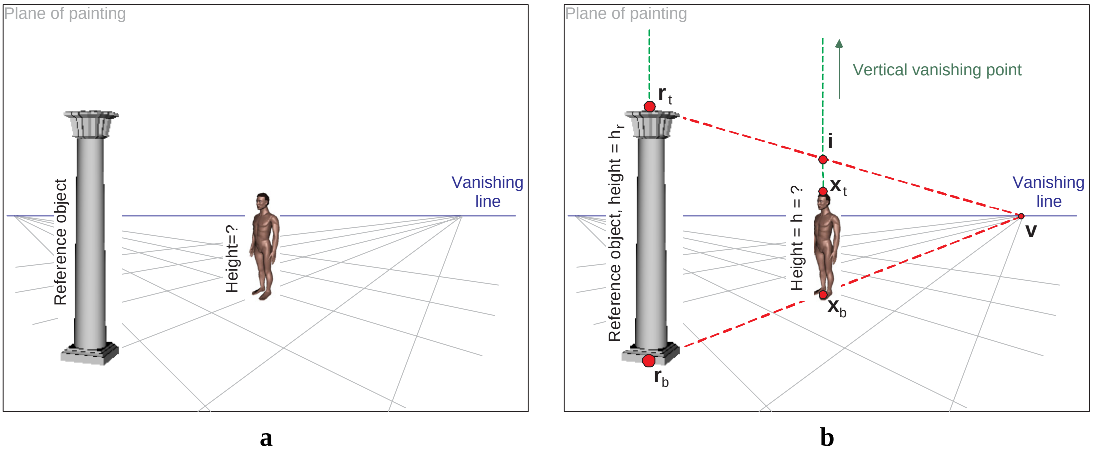


Figure 1: *Measuring heights in paintings.* **(a)** *We wish to compute the height of the human figure relative to the height of the column. The vanishing line of the image is supposed to have been computed (shown in blue).* **(b)** *The ratio between the height of the man and that of the column is given by* $\frac{h}{h_r} = d(\mathbf{x}_t, \mathbf{x}_b)/d(\mathbf{i}, \mathbf{x}_b)$. *See text for details.*

Finally, the ratio between the height of the man and the reference height is simply computed as a ratio between measurable image quantities as

$$\frac{h}{h_r} = \frac{d(\mathbf{x}_t, \mathbf{x}_b)}{d(\mathbf{i}, \mathbf{x}_b)}$$

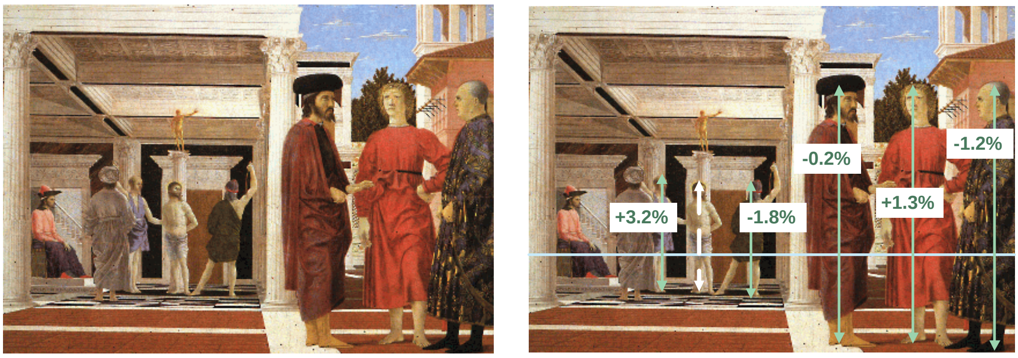
In the case of photographs of real objects the reference height h_r may be known or can be measured in situ and, therefore, the height of the people in the photo can be computed in absolute terms. When, as in the case of most paintings, the reference height is not known we can only compute the ratio $\frac{h}{h_r}$; i.e. we compute the height of people relative to a chosen unitary reference height.

Notice that in this case we have assumed the vanishing point for the vertical direction to be at infinity (all the verticals are parallel to each other in the image) and a horizontal vanishing line of the ground plane. This is a simplified version of our general, algebraic algorithm which can deal with finite vertical vanishing point and a ground-plane vanishing line in any orientation (see [3] for details).

Examples of the application of such technique are presented in the following section.

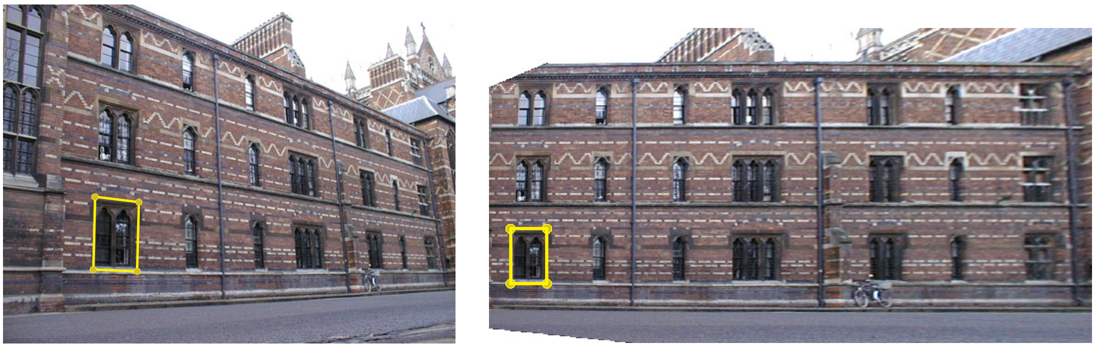
Relative heights in *The Flagellation*. *Flagellation* (in fig. 2a) by Piero della Francesca, is one of the most studied paintings from the Italian Renaissance period. It is a masterpiece of perspective technique. The "obsessive" correctness of its geometry makes it one of the most mathematically rewarding paintings for detailed analysis. In the past, art historians have "dissected" the painting with different laborious techniques, most of them manual [17], with the aim of understanding more about the artist's perspectival constructions.

The metrology algorithm described above has been applied in figure 2b to compute the heights of the people in the painting. Due to the lack of an absolute reference, the heights have been computed relative to a chosen unit reference, i.e. the height of Christ. Therefore, height measurements are expressed as percentage variations from the height of Christ. At a first glance it is not easy to say whether the figures in the background are consistent with the ones in the foreground. But our computer technique has given us the answer: despite little variations the measurements are all quite close to each other, thus confirming the extreme accuracy and care in details no less than in the overall space for which Piero della Francesca has become famed [10].



a b

Figure 2: *Comparing heights of people in a Renaissance painting.* (**a**) *The original painting:* Flagellation *(approx. 1453), 58.4 × 81.5 cm, by Piero della Francesca (1416–92), Galleria Nazionale delle Marche, Urbino, Italia. Courtesy of Ministero per i Beni e le Attivita' Culturali.* (**b**) *The heights of people in the foreground and in the background have been measured relative to the height of Christ. They are expressed in percentage difference.*



a b

Figure 3: *Rectifying planar structures in photographs.* (**a**) *A photograph of the Keble College in Oxford, UK, with four selected points superimposed. The selected points are used to compute the homography between the scene plane and the image plane.* (**b**) *The automatically rectified image. The wall appears to be viewed from a front-on position. See text for details.*

3 Analysis of patterns and shapes

Art historians are often interested in analysing details of varying sizes in certain paintings, such as a complex tile pattern, the shape of an arch or the plan of a vault. The manual techniques they use aim at inverting the geometrical process that the artist employed to construct the geometry of that pattern. Computer techniques can help achieve this "geometric inversion" in a faster , more efficient and comprehensive way.

This section deals, primarily, with the generation of new views of planar patterns. This can be achieved by using plane-to-plane homographies. A plane-to-plane homography is a bijective projective transformation mapping pairs of points between planes [6]. If the homography between a plane in the scene and the plane of the image (the retina or the canvas) is known, then the image of the planar surface can be rectified into a front-on view.

The world-to-image homography can be computed simply by knowing the relative position of four (at least) points on the scene plane and their corresponding positions in the image. Figure 3 shows an example. Figure 3a is a photograph of a flat wall of a building. The four corners of the window at the bottom left corner of the image have been selected and the homography between the plane of the wall and that of the photograph has been computed. The computed homography maps the selected four image points to a rectangle with the same aspect ratio as the window. Thanks to the homography the original image has been warped via software into the rectified image shown in fig. 3b.

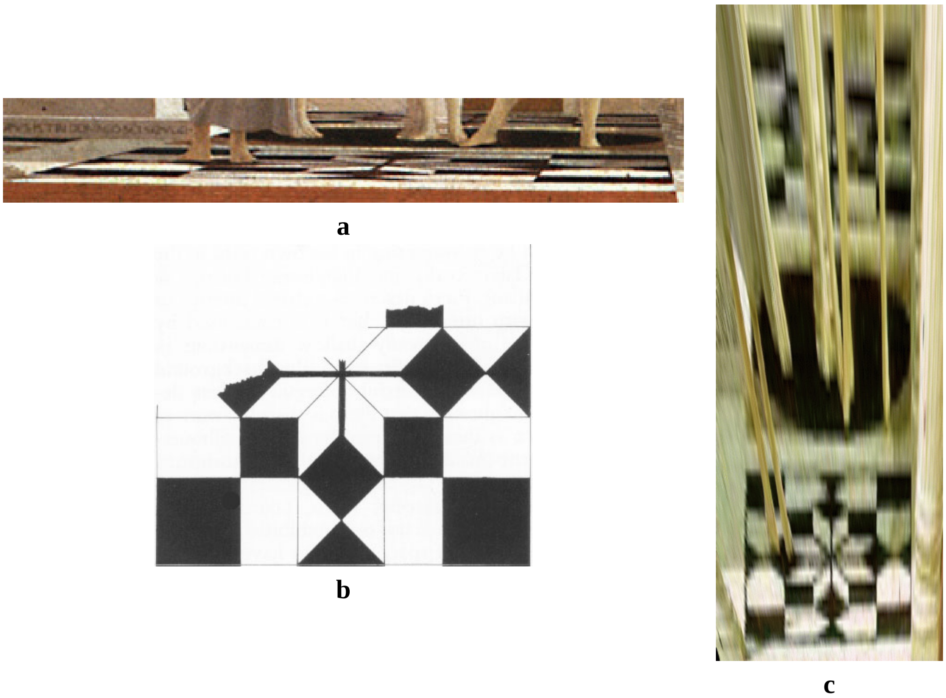


Figure 4: *Analysing a floor pattern, by hand and by computer. (***a***) An image of the floor pattern cropped from fig. 2a. (***b***) Manual rectification of the pattern achieved by Kemp [17]. (***c***) Automatic rectification by computer. The rectified image has been obtained by applying our rectification algorithm to the image of the painting directly. The vertical streaks are the projectively distorted legs.*

Therefore, a new view of the wall has been generated as if it was looked at from a front-on position.

Notice that no knowledge of camera pose or internal parameters (e.g. focal length, type of lens) was necessary. Furhtermore, the quality of the rectification depends on the accuracy in selecting the four basis points. In the example in fig. 3 those were selected by intersecting pairs of straight Canny edges (see [3], pag.33), thus achieving subpixel localization accuracy. The same rectification of slanted planar structures can be performed in perspective paintings as shown in the following section.

3.1 Rectifying the floor of the *Flagellation*

Piero della Francesca's *Flagellation* (fig. 2a) shows, on the left hand side, an interesting black and white floor pattern viewed at a grazing angle. Kemp in [17, 18] has manually analysed the shape of the pattern and demonstrated that it follows the "square-root-of-two" rule. Figure 4b shows the manually rectified image of the floor pattern patiently achieved by Kemp on the basis of a full-sized reproduction. The rectified pattern seems to be observed from above rather than at a grazing angle as in the original painting.

Figure 4c, instead, shows the rectification achieved by applying a homography transformation as described above. In this case the four vertices of the black and white pattern have been selected as the base points for the computation of the homography. We have imposed those to be arranged as a perfect square. Notice the similarity between the computer- and the manually-rectified patterns (fig 4c and fig 4b respectively).

Some of the advantages of the computer rectification are: speed of execution, accuracy and the fact that the rectified image retains the visual characteristics of the original painting. In fig. 4c the original light and color of the pattern have been preserved. Notice, for instance, the interesting shadow line cutting the bottom pattern horizontally. Furthermore, figure 4c shows that two identical instances of the black and white pattern exist in the painting, one before and one behind the central dark circle on which Christ is standing. The farther instance of the pattern is hard to discern by eye in the original painting, while it becomes evident in the rectified view (*cf.* top of fig. 4c).

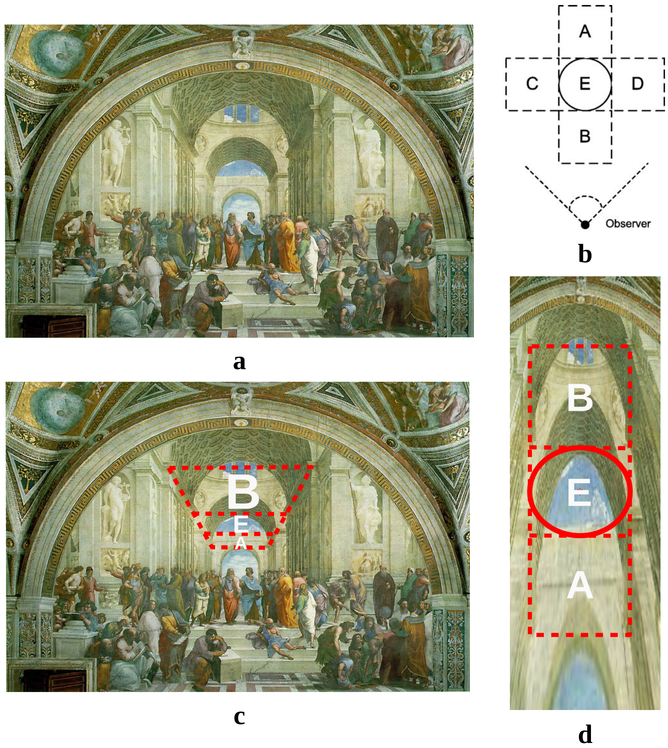


Figure 5: *Analysing the shape of the dome in* The School of Athens*.* (**a**) *The original fresco:* The School of Athens *(1510-11), fresco by Raphael (Raffaello Sanzio, 1483-1520). Stanza della Segnatura, Vatican Museums, Rome.* ©*Vatican Museums.* (**b**) *The plan of the building appears to be a Greek Cross, a cross with two equal-length arms.* (**c**) *The rectangular base of the vault is highlighted in red. In the central part the red ellipse represents the base of the dome.* (**d**) *The base of the vault has been rectified by using a homography transformation. The top and base quadrilaterals (B and A) have been assumed to be square. The central quadrilateral (E) turns out to be a rectangle and not a square as expected. Therefore, the inscribed base of the dome is an ellipse (an ovulum) rather than a circle.*

3.2 The Shape of the Dome in Raphael's *School of Athens*

The previous two sections have demonstrated how by applying homography transformations on portions of paintings it is possible to create new and compelling views of interesting patterns that might warrant analysis, both in the their own right and to ascertain the lengths to which the artist has gone in constructing the painted space. This section shows that homographies may be used also to analyse the shape of non-planar patterns such as domes and vaults.

Raphael's *School of Athens* fresco (in fig. 5a) owes much of its fame to the great airy space of harmonious architecture inhabited by the renowned philosophers of antiquity, lead by Plato and Aristotle. The lucid, classical forms of the building exude an air of rationality and simplicity such that we assume it is based on the geometry of regular rectilinear and circular figures.

We automatically assume that the artist has represented the "school" as a building with the "Greek cross" plan that was so admired in the Renaissance. Since the Greek cross is characterized by two arms of equal length (*cf.* fig. 5b) the quadrilaterals composing the shape of the base (labelled with letters from A to E in figure) should be square in the three-dimensional scene. This would also imply that the central dome, whose base is inscribed in the central quadrilateral (denoted with the letter E), must have a circular base. Our analysis proves that these assumptions are not consistent with the geometry of the painting and that the base of the dome is elliptical (an ovulum), rather than circular. Alternatively, if we assume that the base of the dome is circular, the arms of the cross would consequently need to be longer than they are broad. Either way, the principles of classical design have been subverted.

A planar homography is employed here to rectify the base of the visible vault (the arm of the cross defined by the quadrilaterals A, E and B) into a rectangle, based on the assumption that the two quadrilaterals A and B (in fig. 5c) are perfect squares (in the three-dimensional scene). In this case the homography is defined by the vertices of the three quadrilaterals selected in the original painting (eight points). The computer-rectified image is shown in fig. 5d. Notice that since the vault lies above the plane of its base it gets warped in an unexpected way. In fig. 5d the base of the vault, delineated by the dashed red lines, is rectangular and the quadrilaterals A and B are perfectly square, by construction. Notice that just imposing that one of the two quadrilaterals, A or B is square implies that the other one is square too. The same does not apply to the central quadrilateral, E.

In order to establish the shape of the central quadrilateral (E) it is sufficient to compute the ratio between its height and its width.

$$Ratio_E = \frac{height_E}{width_E} \tag{1}$$

Therefore, $Ratio_E = 1$ for a perfect square and $Ratio_E \neq 1$ for a rectangular-shaped central quadrilateral.

In our experiment (run on the rectified fig. 5d) we measured $Ratio_E = 0.877$ which corresponds to a relative error

$$Err = |Ratio_E - 1.0| = 12.33\% \tag{2}$$

The error $Err >> 0$ confirms that the quadrilateral E is not a square but a rectangle. Since $Ratio_E < 1$ its base is larger than its height. Therefore, the inscribed curve is *not* a circle but an ellipse, with the ratio between the two diameters being $Ratio_E$.

The careful reader may be concerned about the accuracy of these results. In fact, the accuracy of (1) depends on how the eight corners of the three visible quadrilaterals A, E and B are selected on the painting. To achieve a high degree of accuracy we chose to select those points by intersecting pairs of straight Canny edges thus achieving a subpixel level of accuracy.

Furthermore, in order to get a better feeling of the accuracy of the results the experiment was repeated for different sets of input vertices obtained by displacing their location by a few pixels. The relative error in (2) was found to be consistently larger than 10%, thus confirming the elliptical nature of the central dome. These results were also reproduced by a careful (and lengthy) manual analysis, thus confirming the potentiality of our computer technique. A more complete uncertainty analysis may be performed by employing the statistical techniques described in [6].

Since we know from the cartoon (full scale-drawing) for the lower part of the fresco (preserved in the Biblioteca Ambrosiana in Milan, Italy) and the lines incised by the artist in the damp plaster of the wall that Raphael went to enormous trouble to construct the perspective of the "school", this deviation is unlikely to have been causal. The move away from strict regularity is of the kind that artists habitually make when they are undertaking the actual painting and trying to make things "look right" subjectively rather than conforming meticulously to the rules of perspective [2]. It may be that Raphael wanted to "pull" the space under the dome closer to the central figures of Plato and Aristotle by moving the rear arch of the crossing forward, thus enhancing the paradoxical effect that they are standing under the dome.

4 Exploring the paintings third dimension

In the previous sections we have presented algorithms for: evaluating the internal consistency of the geometry of a painting with respect to linear perspective; rectifying and measuring planar surfaces; estimating distances from planar surfaces (e.g. heights of people). By combining all the above techniques together it is possible to create complete three-dimensional models of the painted scenes. Details about the reconstruction algorithms can be found in [3].

This section presents one of the three-dimensional models that have been constructed and explores the possible ambiguities in reconstruction that may arise[3]. More importantly, three-dimensional reconstruction is used also as a tool to detect and magnify possible imperfections in the geometry of the painting.

[2] A classical example is Andrea Mantegna's *The Lamentation over the Dead Christ* (1490, Pinacoteca di Brera, Milan). There the size of Christ's feet has been reduced to make them "look right".

[3] The constructed 3D VRML models can be viewed at `www.robots.ox.ac.uk/~vgg/projects/SingleView/`

Figure 6: *Three-dimensional reconstruction of Masaccio's* Trinity. *(***a***) The original fresco:* The Trinity with the Virgin and St John *(approx. 1426),* 667 × 317 *cm, by Masaccio (Tommaso di Ser Giovanni Guidi, 1401–1428), Museo di Santa Maria Novella, Florence, Italia. Courtesy of the Musei Comunali di Firenze. The reconstruction algorithm is described in [3]. (***b-e***) Different views of the reconstructed three-dimensional model of the chapel in the Florentine fresco.*

4.1 The virtual *Trinity*

The church of Santa Maria Novella, in Florence boasts one of Masaccio's best known frescoes, *The Trinity* (in fig. 6a). The fresco is the first fully-developed perspectival painting from the Renaissance that uses geometry to set up an illusion in relation to the spectator's viewpoint . Masaccio was the first painter to understand and apply Brunneleschi's newly discovered rules of linear perspective, and it has sometimes been supposed that Brunelleschi was involved in its design. This novel way of "creating a three-dimensional illusion on a flat surface" was applied by Masaccio in a series of works, including the Brancacci Chapel in Florence, painted in the span of a few years before his early death in 1428 at the age of 27.

The *Trinity* has been repeatedly analysed using traditional techniques, but no consensus has been achieved. There have been unresolved disputes about the position of the figures, cross and platform in space, given the lack of a visible floor for the "chapel". More importantly, it has become apparent that analyses starting with the assumption that the coffers (or the *entablaturates* at the top of the nearest capitals of the columns) are as deep as they are wide result in a different format from those that start with the premise that the plan of the chapel is square.

Single-view reconstruction algorithms have been applied to an electronic image of the fresco to help art historians reach a consensus over those debated disputes. Views of the resulting three-dimensional model are shown in figs. 6b–e. The model geometry has been computed and stored in a VRML format which can be displayed by any of the many existing browsers[4].

The portions of the vault which are occluded by the capitals and by the head of God have been left blank and are shown in black in the figures. Although such regions could have been filled in as described in [5], here we have chosen not to do so, in order to highlight the areas which are not visible from the original viewing position.

---

4 A number of free VRML browsers are freely available on the Internet, e.g. `www.cai.com/cosmo/` `www.parallelgraphics.com/products/cortona`, `www.glview.com/`

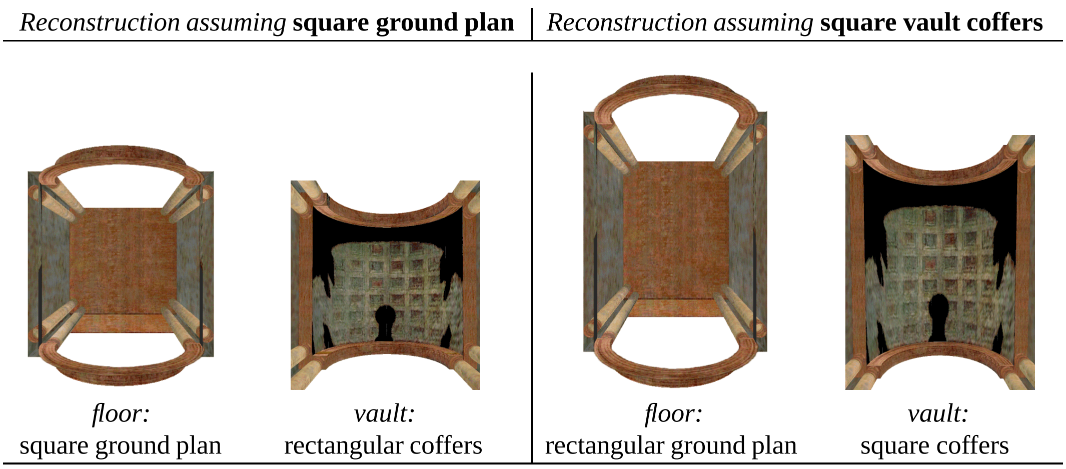


Figure 7: *Ambiguity in reconstructing the depth of the chapel in Masaccio's* Trinity. *Comparing two possible reconstructions from an infinite set of plausible ones. (***Left***) Assuming a square ground plan leads to rectangular vault coffers and (***Right***) Assuming square vault coffers leads to a rectangular ground plan, thus demonstrating that ground plan and coffers cannot be square at the same time.*

Ambiguity in recovering depth. Since one image alone is used and no scene metric information is known (the "chapel" is not real), an ambiguity arises in the reconstruction: it is not possible to uniquely recover the depth of the chapel without making some assumptions about the geometry of the scene.

Two plausible assumptions may be made: either the coffers on the vault of the chapel are square or the floor is square. The reconstruction shown in fig. 6b-e has been achieved by assuming that the vault coffers are square.

Just by looking at the painting one may think that the two assumptions are consistent with each other. Below we demonstrate that that is not the case, the two assumptions cannot coexist, i.e. square coffers imply a rectangular ground plan and vice-versa. Here the two models stemming from the two different assumptions have been computed by means of single-view reconstruction algorithms. Once the first model was constructed, the second one was obtained by applying a simple "affine transformation", a scaling in the direction orthogonal to the plane of the fresco. The advantage in terms of speed and accuracy over manual techniques is blatant.

The image of the chapel floor and that of the vault pattern shown in fig.7for both cases demonstrate that the square-ground-plan assumption yields rectangular coffers and the square-coffers assumption yields a rectangular ground plan.

Because of the lack of visible floor it is not possible to uniquely locate the base of the cross and the human figures. For visualization purposes, in both models we have chosen the cross to be located at the centre of the chapel and the figures of the Virgin and St John at one quarter of the chapel's depth. Once the position of cross and figures has been fixed on the floor their height has been computed consistently, as described in section 2.1.

The number of reconstructions consistent with the original painting is infinite. In fact, different choices of the coffers or ground plan aspect ratios yield different consistent three dimensional models. Only two assumptions (square coffers and square ground-plan), from this infinite set, have been analysed here. Those seem to be more likely than others, for the following reasons:

- Having a square ground plan seems to be the natural choice from a design point of view. Probably, the artist started the design of the chapel by laying down the square foundations, then working out the heights of the columns and finally the rest of the composition.

- The second assumption, that of square coffers, seems to be more likely from a perceptual point of view. The vault coffers are very visible and it makes sense for the observer to subconscially assume them to be square and regularly spaced.

At this point new questions arise: Which architectonical structure wanted the artist to convey? If he had started by laying down a square base, why would he choose rectangular-shaped coffers? Was he aware of the depth ambiguity? Was it done on purpose?

Without exploring the answers in detail here, we suspect that Masaccio began, as most designers would, with the overall shape, and then fitted in the details to look good, and that when he found that his earlier decisions had resulted in coffers that were not quite square (if he noticed!), he decided that they would effectively look square anyway . In the final analysis, visual effect takes over from absolute accuracy. We know, for instance, that Masaccio has intended the circumferential width of the lowest rows of coffers to the left and right of Christ's cross, presumably to avoid visually awkward conjunctions in that area of the painting.

Whatever the reason for Masaccio's ambiguity, the computer analysis performed here has allowed us to investigate both assumptions rigorously, by efficiently building both models, visualizing them interactively and analysing the shape of vault and base in three-dimensions.

A comparison with traditional CAD techniques. An earlier three-dimensional reconstruction of the *Trinity* can be found in [8]. The three-dimensional model constructed in that work was realized by using Computer Aided Design tools. The geometric shape of the scene was constructed by making use of the painting geometry and a number of inevitable yet arbitrary assumptions which find little justification in the original painting.

With our technique, instead, the three-dimensional geometry is extracted directly and exclusively from the image plane and only the objects for which enough geometric information is present in the painting are reconstructed.

In general, architectonical elements in the paintings contain much greater geometrical information (e.g. parallelism and orthogonality of lines and regularity of patterns) than smooth surfaces such as human bodies. The algorithm presented here is capable of constructing architectonical elements up to a very limited number of degrees of freedom. Furthermore, the effect of the remaining ambiguities can be rigorously analysed. In fact, our mathematical tools have enabled us to parametrize the entire infinite family of three-dimensional models which are geometrically consistent (producing the same image from the same vantage position) and satisfy all the assumptions made (as in the case of Masaccio's *Trinity*).

Instead, the human figures are characterized by much less purely geometrical information. Therefore, their reconstruction is subject to a much greater number of degrees of freedom [23], thus leaving us with greater choices about the possible shapes of the human bodies in the painting. Because of this fundamental uncertainty, in this work we have chosen to approximate human bodies with flat cut-outs. It is our intention to investigate the use of shading, texture [7] and contours to try and constrain the reconstruction of smooth surfaces and, possibly, come up with parametric models of clothed human figures of manageable size.

Generally, traditional CAD techniques do not allow this level of reasoning which so profoundly characterizes our methods. The unavoidable assumptions made during the reconstruction process are made explicit in our system and the effect of removing or adjusting them is analysed by efficiently assessing the consequences directly on the reconstructed three-dimensional model.

Finally, not only the geometry, but also the appearance (the texture) of the final model is extracted directly from the image of the fresco, thus achieving a superior level of photorealism, as supposed to the often artificial look of synthetic CAD models in Art History.

5 Conclusion

This paper has presented a number of computer algorithms for: assessing the internal consistency of the geometry in a painting and its conformity to the rules of linear perspective; generating new views of patterns of interest; reconstructing occluded areas of the painting; measuring and comparing object sizes; constructing complete three-dimensional models from paintings; exploring, in a systematic way, possible ambiguities in reconstruction, and, finally, assessing the accuracy of the reconstructed three-dimensional geometry.

The presented algorithms, heavily based on the use of algebraic projective geometry, are rigorous and easy to use. Art historians may use them for a more systematic and efficient analysis of works of art that employ systematic means for the construction of space.

The validity of our approach has been demonstrated, whenever possible, by comparing the computer-generated results to those obtained by the traditional manual procedures employed by art historians.

Furthermore, this paper has proven that computer tools, built on top of a strong projective geometry basis, can answer questions like: How correct is the perspective in any painting posing strong perspectival clues, such as Signorelli's *Circumcision*? [5] What does the pattern on the floor in Domenico Veneziano's *St Lucy Altarpiece* look like?

The ability to turn flat paintings into interactive three-dimensional models opens up a new and exciting way of experiencing art with new levels of vividness and historical awareness. Inverting the painting process and navigating inside three-dimensional scenes may be used for teaching purposes. Art students and aficionados may be able to understand better the power of linear perspective by interacting "hands on" with three-dimensional objects rather than by just looking at the painting. Art historians will be able better to judge the levels of concern of artists with perspectival accuracy and to analyse possible reasons for their departure from strict obedience to geometrical rule. The relationship between the "ideal" viewpoints and those available to the spectator in actual locations, such as S. Maria Novella, the site of Masaccio's Trinity, may be easily explored by using these tools. Another interesting application is to set the paintings in virtual reconstructions of their original locations (if known).

A new, XML-based visualization tool has been presented where the user can navigate inside a virtual museum and, seamlessly fly from the museum to the paintings 3D scenes and back.

Our techniques are currently being employed in the *Empire of the Eye* project under design for the National Gallery in Washington[6]. *Empire of the Eye* intends to build upon and extend *Masters of Illusions* [16]. One of the project's aims is that of interactively teaching about the power of linear perspective in conveying the three-dimensional illusion on a flat surface. The tools described in this paper prove invaluable for constructing interactive three-dimensional environments from historical paintings.

Future work. The focus of this paper has been the geometry of paintings, therefore, our examples have considered works from the Italian Renaissance and seventeenth-century Holland, two of the eras in which the rules of linear perspective were most rigorously applied. The potential of this method for viewing of paintings in situ has already been suggested.

Currently our interests focus on a very different aspect of painting, i.e. the management of light and shading. We are investigating the use of light, shading and colours in paintings and exploring possible ways to model those effects in a rigorous and systematic way [20].

The relationship between the distribution and light, shade and colour in paintings and the scientific notions of light available to the artists (from Aristotle to Newton and later) is a meter of complex historical debate [17] it should be possible to take some of the artists who adopted a consciously experimental approach to light and colour, such as the Impressionists and Neo-impressionists (like Seurat) and to place their creations in dialogue with the techniques of computer vision and computer graphics.

The period that best represents the theories of light and colour is the French Impressionism. Artists like Monet and van Gogh were amongst the first to experiment with the scientific theories of light composition and mixing of colours discovered in the seventeenth century by Isaac Newton.

Furthermore, some recent work of ours has dealt with the analysis of the accuracy of painted mirrors [4] and the implications on Hockney's recent theories [14].

Discussion. When looking at photographs of a scene, visual cues such as converging straight lines, shading effects, receding regular patterns, shadows and specularities are processed by our brain to retrieve consistent information about the surrounding environment.

The same visual cues have over the years come to be employed by artists in their paintings. However, works of art are rarely designed to conform precisely to a set of optical rules and are not presented as scientific theorems. Thus, the visual signals might not be consistent with each other. However, as a large amount of psychophysics research shows [1, 2], our brain is capable of neglecting conflicting cues, and slightly inaccurate perspective or shading may

[5] *Circumcision*, by Luca Signorelli (1445-1461), National Gallery, London

[6] `www.nga.gov`

still convey the desired three-dimensional illusion. Also, it may be apparent that breaking the rules creates a forceful artistic effect.

A number of interesting questions arise: Which perceptual cues are most important to the three-dimensional illusion? To what extent do humans forgive wrong cues? How much does each artist make use of visual cues? Which ones? How about abstract or cubist paintings ? These points may lead the way to further interesting psychophysical speculations in which science and art can play a fundamental role.

It is the authors opinion that computer graphics, computer vision, art history and perceptual psychology and neurology, each of which are well-distinguished disciplines with their own aims and motivations may learn from and be enriched by the others. Furthermore, the tools developed in one area may be transferable productively to another. The present paper is but a step in that process.

Acknowledgements. The authors would like to thank Ian Reid and David Liebowitz from the University of Oxford, Luc van Gool from the Katholieke Universiteit of Leuven and Steve Seitz from the University of Washington for their help and suggestions and Matthew Uyttendaele, Rick Szeliski and P. Anandan from Microsoft Research for useful and inspiring discussions and for providing the XML-based visualization tool. Also thanks are due to the collaborators of the *Empire of the Eye* project: Joseph Krakora, Jay Levenson and Curtis Wong.

References

[1] A. Blake and H. Bultoff. Does the brain know the physics of specular reflection. *Nature*, 343(6254):165–168, 1990.

[2] H.H. Bulthoff and H.A. Mallot. Interaction of different modules in depth perception. In *Proc. 1st International Conference on Computer Vision, London*, pages 295–305, 1987.

[3] A. Criminisi. *Accurate Visual Metrology from Single and Multiple Uncalibrated Images*. Distinguished Dissertation Series. Springer-Verlag London Ltd., September 2001. ISBN: 1852334681.

[4] A. Criminisi, M. Kemp, and S. B. Kang. Reflections of reality in Jan van Eyck and Robert Campin. In *Measuring Art: A Scientific Revolution in Art History*, Paris, May 2003.

[5] A. Criminisi, M. Kemp, and A. Zisserman. Bringing pictorial space to life: computer techniques for the analysis of paintings. In *on-line Proc. Computers and the History of Art*, British Academy, London, Nov. 2002.

[6] A. Criminisi, I. Reid, and A. Zisserman. A plane measuring device. *Image and Vision Computing*, 17(8):625–634, 1999.

[7] A. Criminisi and A. Zisserman. Shape from texture: homogeneity revisited. In *Proc. 11th British Machine Vision Conference, Bristol*, pages 82–91, UK, September 2000.

[8] M. De Mey. Perspektief in 3-D. *Gent Universiteit*, 8ste jaargang(3):14–17, December 1993.

[9] P. E. Debevec, C. J. Taylor, and J. Malik. Modeling and rendering architecture from photographs: A hybrid geometry- and image- based approach. In *Proceedings, ACM SIGGRAPH*, pages 11–20, 1996.

[10] Piero della Francesca. *De Prospectiva Pingendi*. Firenze, Italy, 1474. Reproduced by ed. Sansoni (1942), Edizione Critica.

[11] O. D. Faugeras. *Three-Dimensional Computer Vision: a Geometric Viewpoint*. MIT Press, 1993.

[12] J. V. Field. *The Invention of Infinity, mathematics and arts in the Renaissance*. Oxford University Press, 1997.

[13] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521623049, 2000.

[14] D. Hockney. *Secret Knowledge: Rediscovering the Lost Techniques of the Old Masters*. Viking Press, October 2001.

[15] Y. Horry, K. Anjyo, and K. Arai. Tour into the picture: Using a spidery mesh interface to make animation from a single image. In *Proceedings of the ACM SIGGRAPH Conference on Computer Graphics*, pages 225–232, 1997.

[16] J.Burke. Masters of illusion. 1991.

[17] M. Kemp. *The Science of Art*. Yale University Press, New Haven and London, 1989. ISBN: 0-300-05241-3.

[18] M. Kemp. *Visualizations: the nature book of art and science*. The University of California Press, Berkeley and Los Angeles, California, USA, 2000. ISBN: 0-520-22352-7.

[19] MetaCreations. http://www.metacreations.com/products/canoma/.

[20] S. Nayar and S. Narasimhan. Vision in bad weather. In *Proc. 7th International Conference on Computer Vision, Kerkyra, Greece*, pages 820–827, September 1999.

[21] J. G. Semple and G. T. Kneebone. *Algebraic Projective Geometry*. Oxford University Press, 1979.

[22] B. Taylor. *Linear Perspective*. London, 1715.

[23] L. Zhang, G. Dugas-Phocion, J.-S. Samson, and S. Seitz. Single view modeling from free-form scenes. In *Proceedings of the Conference on Computer Vision and Pattern Recognition, Kauai Marriott, Hawaii*, December 2001.