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

antcrim@microsoft.com

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.

Contents

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 [22] – 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* [14, 33, 36]. 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 [13, 15, 10]. In particular, the work on single-view metrology [6, 18, 24] 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

[1] The multiple-view techniques developed in the Computer Vision and Computer Graphics communities such as [10] are not applicable to accuratelly reconstructing and analysing paintings because of the lack of stereopsis and scene measurements.

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 [22] amongst others. The analysis of a painting should preferably be performed on as large a reproduction as is available, ideally life-size (though this is rarely possible). The advantages of the hand-drawn analyses are:

- 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);
- 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:

- it is all to common to find thick lines drawn on small reproductions of big paintings, with resultant imprecision;
- 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.
- 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 [9]. The advantages of the CAD Reconstructions are:

- they depict spatial features with precision according to the classic rules of linear perspective;
- using techniques for rendering, they produce pictorial effects of light, shade, colour and (to a degree) texture akin to those in the original image;
- 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:

- 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;
- 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:

- 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;
- 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;
- the textures and colours of the original image are retained;
- re-projection can scrutinise errors and allow for their correction;
- degrees of inaccuracy can be estimated systematically across the surface of images;
- fly-throughs and externalised views of the space can be produced with ease;
- regions occluded by objects closer to the viewer can be systematically reconstructed.

The disadvantages are:

- the power of the analysis may be excessive for the quality of the information which the artist entered into the original painting;
- it is applicable with profit only to those paintings which were constructed with sustained attention to perspectival rules;
- 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.

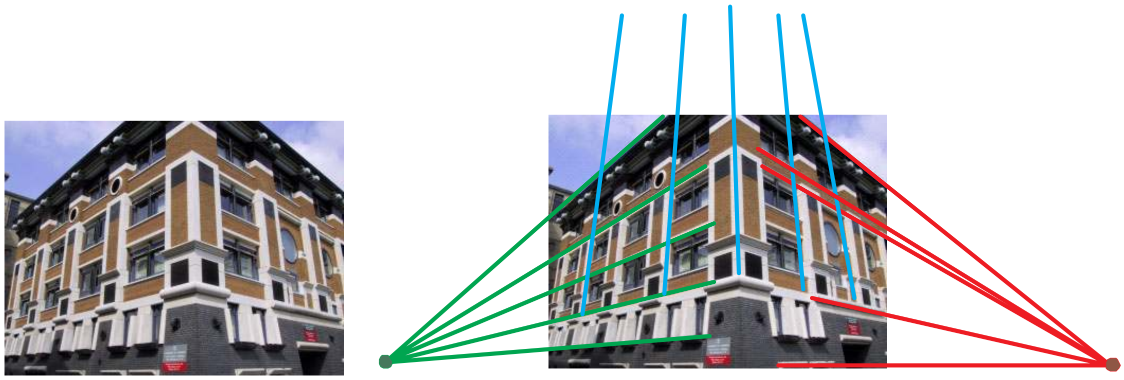
We like to think of the technique described in this paper as a more flexible Computer-Assisted analytical tool which, unlike other CAD applications, allows the user to analyse the geometrical information contained in the painting in a more rigorous way by exploring all the possible reconstruction alternatives in a rigorous way, 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:

- 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;
- evidence from drawings by the artist and comparable artists about the methods they employed to construct perspectival spaces;
- 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, and fill in occluded regions. 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. Finally in section 5 we show our "Virtual Museum", an interactive, three-dimensional



a **b**

Figure 1: *In a photograph parallel scene lines are imaged as converging lines. (***a***) A photograph of a building in Cambridge, UK. (***b***) Three sets of parallel lines (in green, blue and red) are shown. In the image plane all the lines in each set converge into a single point, namely the vanishing point. The vanishing point associated with the set of lines shown in blue falls well outside the image.*

visualization tool which enables us, not only to observe paintings in an interactive three-dimensional environment, but also to virtually "dive" into the painted scene. The virtual museum application is also shown to be useful as a way of measuring the correctness of the three-dimensional reconstruction of a painted scene. Throughout the paper the labourious analytical work of patient art historians is shown side by side with the results originated by applying our fast 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.

2.1 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 (*cf.* fig 1). 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.

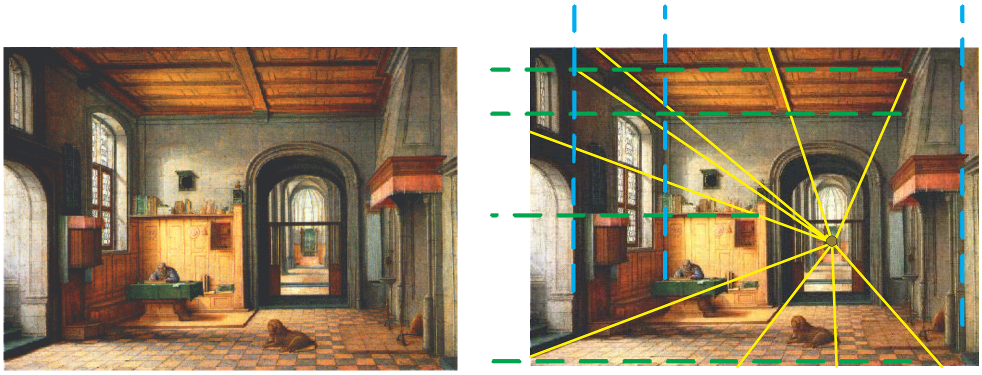
Vanishing points characterize any perspective image and therefore also perspective paintings. Figure 2 shows the vanishing points associated with three orthogonal directions in a painting. A special case is that of a vanishing point at infinity. This occurs when the image lines are all parallel to each other (*cf.* the sets of blue and green lines in fig. 2) and arises from scene lines parallel to the image plane/canvas.

In the past, vanishing points have been used extensively by artists to convey a three-dimensional illusion of the represented scene and also to focus the attention of the observer towards points of interest, e.g. by placing them in strategic locations such as the face or the hands of an important figure[2], or the centre of an open door[3].

The configuration of vanishing points present in fig. 2 is typical of many paintings, i.e. only one vanishing point is finite and close to the centre of the image. Often, the lines emanating from the finite vanishing point are called *orthogonals* because they are orthogonal, in the scene, to the plane of the painting, i.e. the canvas, panel or wall surface. This arrangement of vanishing points arises from a special relationship between the image-plane/canvas and

[2]e.g. in *The Last Supper* by Leonardo da Vinci (1452-1519), Convent of Santa Maria delle Grazie (Refectory), Milan, Italy

[3]e.g. in the *Ideal City*, formerly attributed to Piero della Francesca, National Gallery, Urbino, Italy



a **b**

Figure 2: *Vanishing points in a painting. (***a***)* St Jerome in his Study *(1630),* 40×56.2 *cm, by Hendrick V. Steenwick (1580–1649), private collection (Joseph R. Ritman Collection), Amsterdam, The Netherlands (***b***) The vanishing points associated with the three main orthogonal directions. Notice that two vanishing points (those associated with the horizontal and vertical directions) are at infinity.*

principal scene directions (where two of the principal scene directions are parallel and the third is orthogonal to the canvas). However, in general photographs of a scene this special arrangement will not hold. This is why in the computer vision community lines emanating from a vanishing point are not given any specific name.

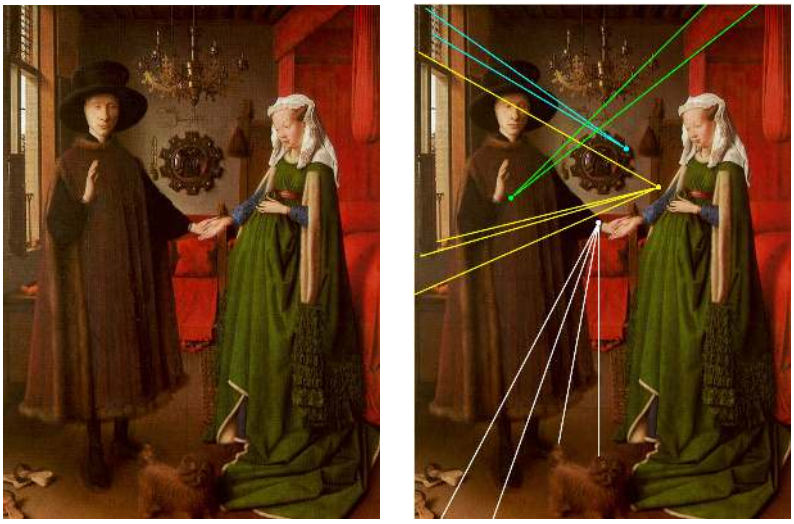
Two or more aligned vanishing points define a *vanishing line*. A good example of a vanishing line is the *horizon*. This is also depicted in many paintings and defines the *eye level* (*cf.* fig.8b). In mathematical terms, a vanishing line is associated with a set of planes parallel to each other in the scene. Therefore, in general, the vanishing line is not necessarily horizontal and can take any orientation on the image plane, its direction being related to the orientation of the corresponding planes in the three-dimensional space.

From the above we can conclude that the geometry represented in paintings is generally simpler than that of most photographs of complex scenes. Since vision techniques such as single-view reconstruction were designed to work on photographs they do work as well on perspectively consistent paintings.

Vanishing points and vanishing lines are amongst the most useful projective entities of an image. A natural way to assess the correctness of a painting geometry is to check whether images of parallel lines do intersect in a single point on the painting. Figure 3 shows the *Arnolfini portrait* by Jan van Eyck. A number of straight lines can be observed: e.g. the slabs on the wooden floor, the edges of the window on the left and the edges of the bed on the right. An obvious assumption to make is that the bottom edge of the window is horizontal in the scene. If that is true then, in the image plane, the window's bottom edge should go through the same vanishing point as the edges of the floor slabs. However, it is easily shown that this is not the case.

Figure 3b shows some of the lines which have been extracted from the painting. These have been computed by: (i) running a standard Canny edge detector [5] and (ii) fitting straight lines to the computed edges. For a correct geometry the white and the yellow lines should converge at the same vanishing point. The fact that this does not happen is a sign of geometric inconsistency and, given David Hockney's theories suggesting that the painting may have been constructed piece by piece with the help of optical devices for some of the more complex forms, the lack of a consistent vanishing point does not come as a surprise. Hockney [17] suggests that the *Arnolfini Portrait* may have been constructed as a collage of images captured from different locations by a camera-like optical tool, perhaps just a concave mirror. Hockney's hypothesis is validated here by the fact that no globally consistent vanishing point exists.

In fact, the position of vanishing points is unaffected (invariant) to camera translation but is extremely sensitive to camera rotation. The method of constructing a painting by means of an optical tool moved around to depict a section of the composition at a time is quite sensitive to this problem. In fact, even the slightest accidental rotation of the optical device employed by the artist in between stages of the composition would cause the lack of global consistency in the vanishing points of the different sections of the painting. This would cause the painting's geometry to be consistent only at a local level.



a **b**

Figure 3: *Analysing the consistency of vanishing points in a painting. (**a**) The original painting:* Giovanni Arnolfini and His Wife Giovanna Cenami (The Arnolfini Portrait) *(1434) by Jan van Eyck (circa 1395-1441). Tempera on wood. The National Gallery, London, UK. (**b**) Some of the computed vanishing points. For a globally consistent geometry all the marked lines (corresponding to lines parallel to each other in space) should intersect in the same vanishing point.*

Interestingly, despite its lack of global geometric consistency the *Arnolfini Portrait* maintains a striking visual coherence. The painting does not look or feel wrong, and the artist was clearly happy with the spatial effects he has created, even if they are not as internally consistent as those of his Italian contemporaries.

Examples of compositions in which the geometry is only locally consistent are abundantly present amongst the many drawings and prints produced by Escher. Escher's compositions, unlike van Eyck's paintings, openly parade their characteristic geometric inconsistency in terms of a series of visual paradoxes. Two centuries earlier, William Hogarth in his frontispiece to John Joshua Kirby, *Dr Brook Taylor's method of perspective made easy in both theory and practice*[4] had already produced an image full of humorous paradoxes as a way of heightening our awareness of the need for internal consistency. As it happens, Hockney's interests in varieties of perspective and the limitations of the orthodox construction is reflected in the painting that he based in 1975 on Hogarth's frontispiece, *Kerby (after Hogarth) Useful Knowledge* (Museum of Modern Art, New York) [16].

2.2 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 (*cf.* fig 4) 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. The traditional way of checking, dating back to Alberti, is to draw a diagonal across one of the tiles, which, if extended, should pass through the diagonally opposite corners of successive tiles. However, this only works if the floor tiles are square or rectangles of constant size.

This section presents a simple algebraic technique to assess the correctness of the geometry of receding regular patterns (e.g. tiled floors, arched porches and decorated vaults). Patterns like the one depicted in fig. 4a were widely used during the Renaissance to convey a fuller three-dimensional sense of space.

Figure 5a shows a painting by the Dutch artist Peter Saenredam, who is well known for his paintings of church interiors. In fig. 5c we have selected two lines which cut the vault longitudinally. Four points (shown in green) corresponding to the middle arch of each vault section have been selected along each line. Assuming that the vault is

[4]published in Ispwitch in 1754

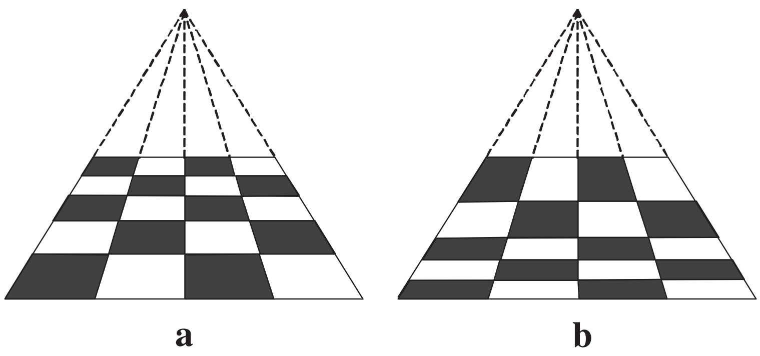


Figure 4: *A drawing of a tiled floor. (***a***) The tiles are perceived to be the same size. (***b***) The tiles are not perceived to be equal size. In this case the rate of diminution of the images of the tiles is not consistent with the assumption that all the tiles have the same area. Despite the fact that in both cases the "orthogonals" correctly converge into a single vanishing point, our brain does not perceive the second as an image of a slanted floor.*

regular (i.e. the different sections in the farther section of the vault have equal sizes) the space between those points in the scene should be constant for both lines. In mathematical notation

$$d(\mathbf{A}_i, \mathbf{B}_i) = d(\mathbf{B}_i, \mathbf{C}_i) = d(\mathbf{C}_i, \mathbf{D}_i) \quad \text{for } i = 1, 2 \tag{1}$$

where $d()$ indicates distance and the capitalized letters indicate points in the three-dimensional scene space, corresponding to the points marked in fig. 5c.

Equation (1) holds only in the three-dimensional scene and not in the image plane. Nevertheless, the distances between the image points $\mathbf{a}_i$, $\mathbf{b}_i$, $\mathbf{c}_i$ and $\mathbf{d}_i$ are constrained and cannot assume any value (*cf.* fig. 4b). In fact, according to the rules of projective geometry, the distances between the four image points along the lines L_1 and L_2 must obey the cross-ratio rule:

$$CrossRatio = \frac{d(\mathbf{a}_i, \mathbf{c}_i) d(\mathbf{b}_i, \mathbf{d}_i)}{d(\mathbf{b}_i, \mathbf{c}_i) d(\mathbf{a}_i, \mathbf{d}_i)} = \frac{4}{3} \tag{2}$$

Equation (2) holds for both lines $\mathbf{L}_1$ and $\mathbf{L}_2$ in the image plane.

From the image measurement we have computed the cross-ratios corresponding to the two lines as:

$$CrossRatio_1 = 1.3765 \qquad CrossRatio_2 = 1.3352$$

In order to assess the accuracy of those measurements it is necessary to compare the above results to $\frac{4}{3}$. By defining the relative error as

$$Err_i = \frac{|CrossRatio_i - \frac{4}{3}|}{\frac{4}{3}}$$

we obtain the following measurements for the two cases

$$Err_1 = 3.2349\% \qquad Err_2 = 0.1413\%$$

Although not identical, in both cases the relative error is very small, thus confirming the correctness of the rate of diminution of the vault sections in Saenredam's painting. On the other hand, we can be reasonably sure that Saenredam had access to a measured survey of St Bavo [30]. Notice that the lines L_1 and L_2 in fig. 5c, in the case of planar surfaces, do not need to go through the vanishing point of the painting, thus adding flexibility to our technique. In this case we have made use of cross-ratios for ease of explanation, but another way of looking at this problem is algebraically, by employing projective mappings known as line-to-line homographies [15]. Line-to-line homographies would add robustness to our technique.

The technique described above may be used for any pattern which presents a regularly repeated structure such as tiled floors, carpets, leaded windows and walls with repeated patterns. In fact given three collinear points which are

a **b** **c**

Figure 5: *Assessing the consistency of receding patterns in paintings. (***a***) The original painting:* Interior of the Church of St Bavo in Harleem *(1648) by Pieter Jansz Saenredam (1597-1665). Oil on panel,* 200 × 140 *cm. National Gallery of Scotland, Edinburgh. (***b***) Detail of (a); the rear vault. (***c***) Two lines have been superimposed and four points have been selected along each line corresponding to the central arch of each vault section. The assumption of a regularly-shaped vault implies that the four points have the same distance from each other in the scene. This translates, in the image plane, into obeying the cross-ratio rule.*

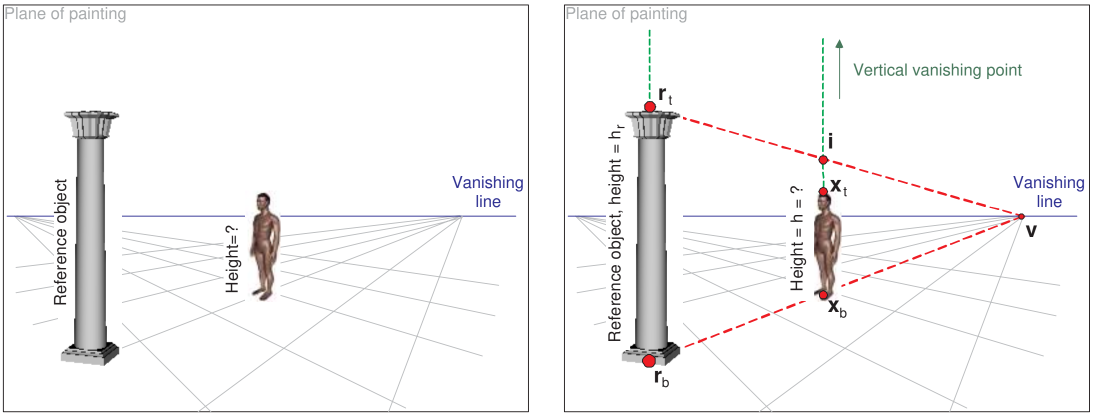
images of three collinear and equally spaced points in the world, then images of all the other equally spaced points on that line can be generated as well as the vanishing point of the line. Similarly, given images of three equally spaced and parallel coplanar lines, then the images of all the other equally spaced parallel lines on that plane can be computed, as well as the vanishing line of the plane [31].

2.3 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. 6 is used to explain how heights of people can be computed directly from perspective images. In fig. 6a 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.



a **b**

Figure 6: *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

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

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 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.

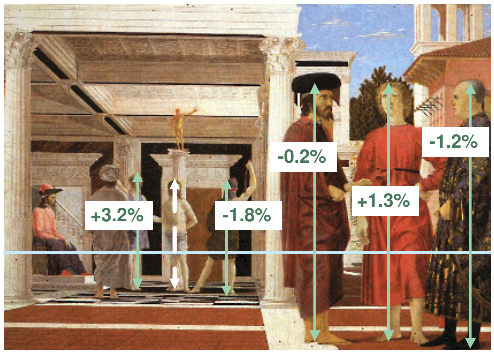
Notice that in this case we have assumed the vanishing point for the vertical direction to be at infinity (all the verticals are parallel to each other in the image) and a horizontal vanishing line of the ground plane. This is a simplified version of our general, algebraic algorithm which can deal with finite vertical vanishing point and a ground-plane vanishing line in any orientation (see [6] for details).

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

Relative heights in *The Flagellation*.

Flagellation (in fig. 7a) 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 [22], with the aim of understanding more about the artist's perspectival constructions.

The metrology algorithms described above have been applied in figure 7b 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 [11].



a **b**

Figure 7: *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.* (**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.*

Relative heights in *The Marriage of the Virgin*.

Figure 8a shows another example of Renaissance art with a developed perspective construction. It was painted by Raphael (Raffaello Sanzio) at the tender age of 21. There, again the human figures are separated into a foreground and a background crowd.

In fig. 8b the heights of some figures have been computed relative to that of the woman on the left. In this case the two measured figures on the background appear to be slightly smaller than the ones on the foreground. In this case it is harder to say whether this is due to a geometric inconsistency or a deliberate desire to exaggerate the perceived depth. In fact, the measurements on the background are much less accurate than the ones on the foreground because of the much larger distance from the observer and the consequent reduced image resolution, and the effect is probably not deliberate. A thorough analysis of the measurement accuracy may be performed by applying the analytical techniques described in [6].

3 Analysis and synthesis 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 [7]. 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 9 shows an example. Figure 9a is a photograph of a flat wall of a building. In fig. 9b 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. 9c. 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 calibration was necessary. Furhtermore, the quality of the rectification depends on the accuracy in selecting the four basis points. In the example in fig. 9 those were selected by intersecting pairs of straight Canny edges (see [6], pag.33), thus achieving subpixel localization accuracy. The same rectification

a **b**

Figure 8: *Comparing heights of people.* (**a**) *The original painting:* The Marriage of the Virgin*, 1504, by Raffaello Sanzio (1483-1520), Brera Museum, Milano, Italy.* (**b**) *Measurements of people heights have been superimposed to the original painting. The computed horizon line is shown in blue.*

of slanted planar structures can be performed in perspective paintings as shown in the following section.

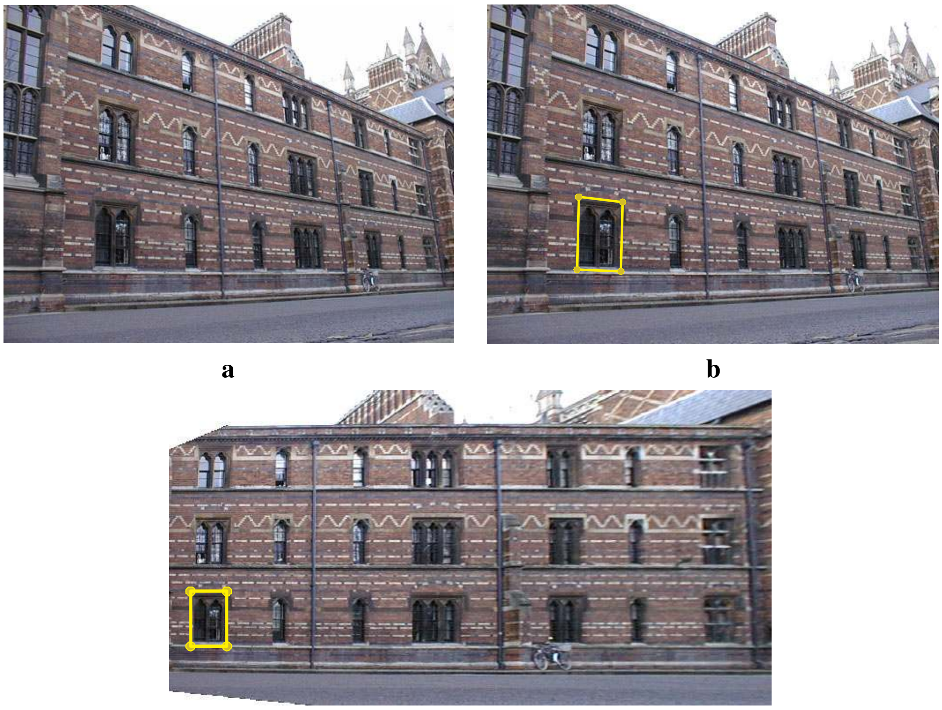
3.1 Rectifying the floor of the *Flagellation*

Piero della Francesca's *Flagellation* (fig. 7a) shows, on the left hand side, an interesting black and white floor pattern viewed at a grazing angle. Kemp in [22, 23] has manually analysed the shape of the pattern and demonstrated that it follows the "square-root-of-two" rule. Figure 10b 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 10c, 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 10c and fig 10b 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. 10c 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 10c 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. 10c). The sharpness of the rectified image decreases going from bottom to top, due to the fact that in the original painting the floor is observed at a grazing angle; therefore, regions of the floor farther away from the observer are characterized by lower image resolution.

Many portions of the floor are occluded by the columns and by the people standing on it. These occlusions show up in the rectified image (fig. 10c) as very thin vertical streaks. These are due to the mapping of the legs of the people onto the plane of the floor from the original viewpoint. They may be thought of as equivalent to shadows cast by the figures on more distant features by a point source of light positioned at the spectator's ideal eye point.



a **b** **c**

Figure 9: *Rectifying planar structures in photographs. (***a***) A photograph of the Keble College in Oxford, UK. (***b***) Original photo with four selected points superimposed. The selected points are used to compute the homography between the scene plane and the image plane. (***c***) The automatically rectified image. The wall appears to be viewed from a front-on position. See text for details.*

3.2 The floor of Veneziano's *St Lucy Altarpiece*

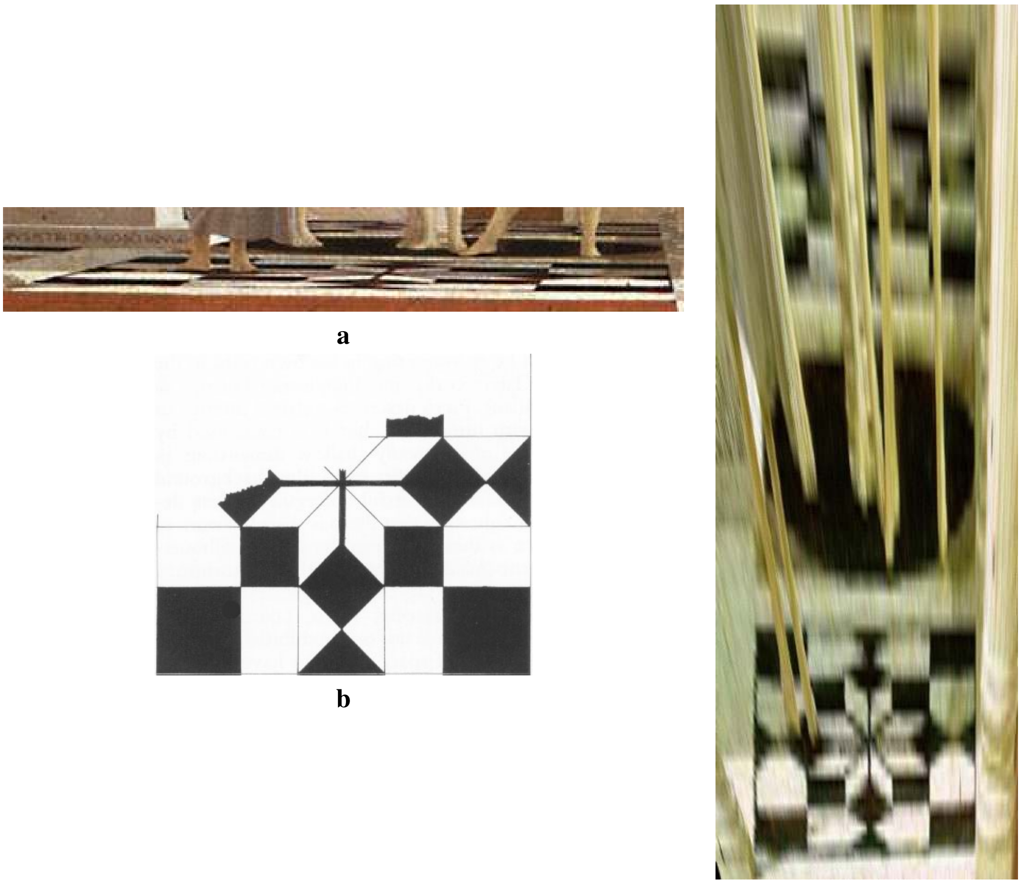
A more complicated pattern can be found on the floor of the *St Lucy Altarpiece* by Domenico Veneziano (fig. 11a). The manually-rectified floor pattern shown in fig. 11b may be found in [22].

As described in the previous section, in the computer-based rectification, four basis points (assumed at the corners of a square) were selected on the floor. The floor-to-image homography was computed and the image automatically rectified. The resulting rectified image is shown in fig. 11c. Once again the similarity between the manually- and the computer-rectified images of the floor is glaring. Working with a higher-resolution image would improve the sharpness of the rectification in fig. 11c. Notice the vertical streaks due to the mapping of the legs of the standing people on the floor. Compared with the manual technique, the automatically rectified image retains the visual characteristics of the painted surface and does not require the kind of geometrical drafting employed by Kemp.

3.3 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. 12a) 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.



a

b

c

Figure 10: *Analysing a floor pattern, by hand and by computer. (***a***) An image of the floor pattern cropped from fig. 7a. (***b***) Manual rectification of the pattern achieved by Kemp [22]. (***c***) Automatic rectification by computer. The rectified image has been obtained by applying our planar warping algorithm to the image of the painting directly. The vertical streaks are the projectively distorted legs.*

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. 12b) 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. This section 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. 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. 12c) 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. 12d. Notice that since the vault lies above the plane of its base it gets warped in an unexpected way. This effect is similar to that "shadow" effect shown in figs 10c and 11c. In fig. 12d 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

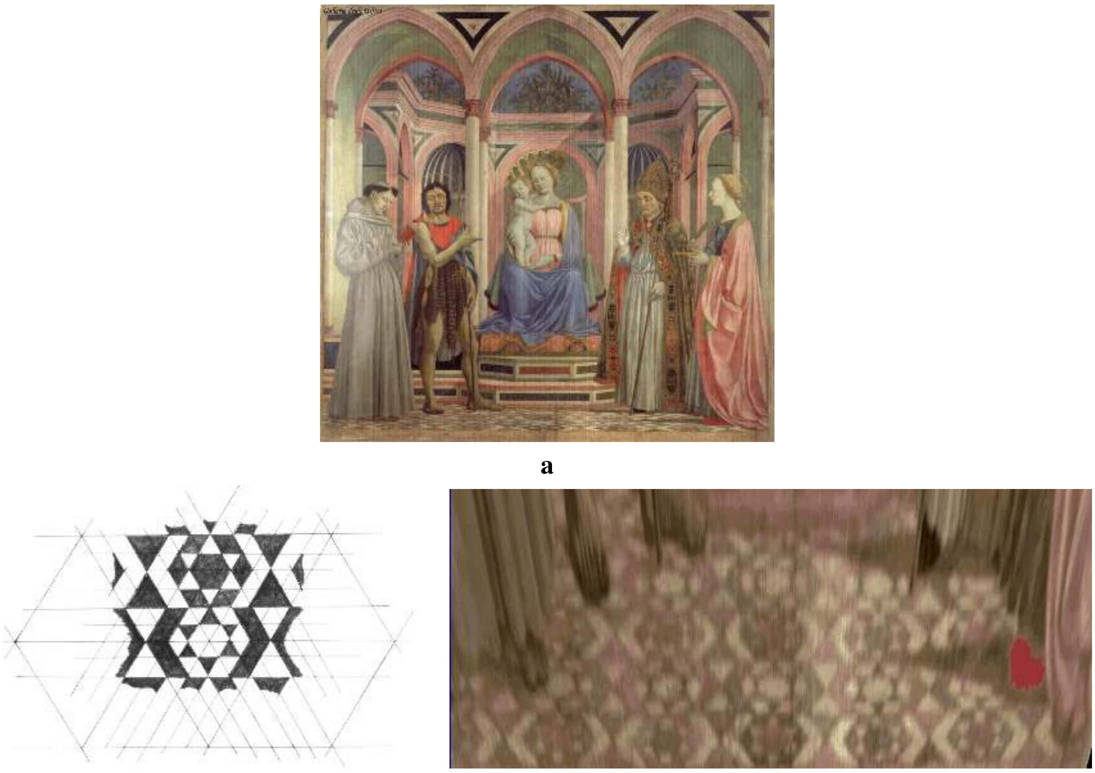


Figure 11: *Analysing a floor pattern, by hand and by computer. (***a***) The original painting:* St Lucy Altarpiece*, 1444, by Domenico Veneziano (1400-1461), Uffizi, Florence, Italy. (***b***) The manual rectification of the floor achieved by Kemp in [22]. (***c***) The computer rectification of the floor obtained by applying our projective geometry-based technique.*

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.

$$Ratio_E = \frac{height_E}{width_E} \tag{3}$$

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. 12d) we measured $Ratio_E = 0.877$ which is a relative error

$$Err = |Ratio_E - 1.0| = 12.33\% \tag{4}$$

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 (3) 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 detected as in section 2.1; 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 (4) 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 [7].

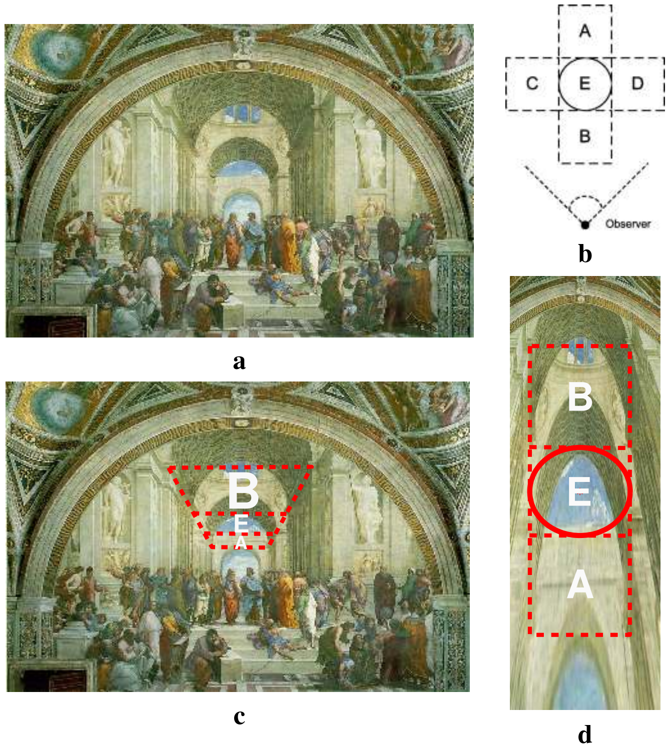


Figure 12: *Analysing the shape of the dome in* The School of Athens*. (**a**) The original fresco:* The School of Athens *(1510-11), fresco by Raphael (Raffaello Sanzio, 1483-1520). Vatican Museums, Stanza della Segnatura, Rome . (**b**) The plan of the building appears to be a Greek Cross, a cross 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.*

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 [5]. 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.

3.4 Completing partially occluded patterns

Section 2.2 has shown a way to analyse the accuracy of the vault pattern in Saenredam's painting of the St Bavo church. In this section a different problem is addressed, i.e. that of recovering the appearance of missing parts of a regular pattern.

In figure 5a, the part of the vault closer to the observer is partially hidden by a large organ on the right hand side. Our aim is to remove the organ from the painting and reconstruct the occluded areas of the vault. This is done by assuming the vault to be symmetric in the three-dimensional space and exploiting this symmetry directly on the plane

[5] 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".

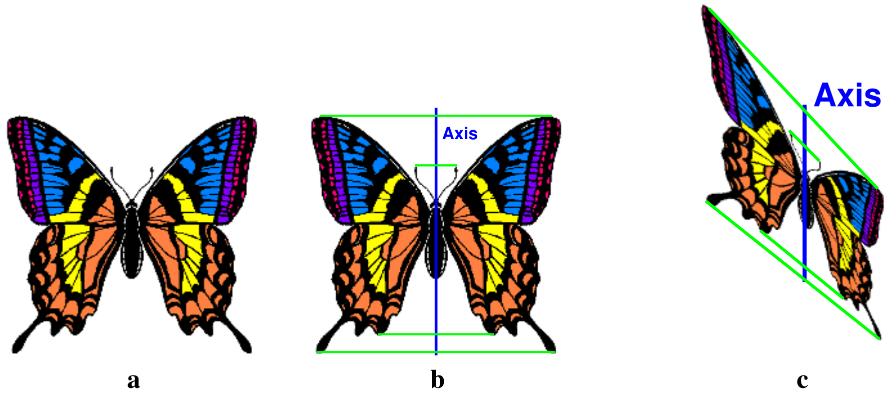


Figure 13: *Bilateral symmetries are imaged as planar harmonic homologies. (***a***) The schematic image of a butterfly. It is characterized by a planar bilateral symmetry. (***b***) The axis of symmetry is drawn in blue. Corresponding points on the two sides are connected by the green segments. Those intersect the axis orthogonally. (***c***) An image of the butterfly taken at a slanted angle is, in general, no longer symmetric. Corresponding points on the two sides are now related by a* planar harmonic homology. *The axis of the homology is drawn in blue. The segments joining corresponding points no longer intersect the axis orthogonally.*

of the painting.

An enlargement of the St Bavo vault is shown in fig. 14a. It is a star-shaped vault composed from eight losange-shaped, concave vaulting sections between supporting ribs, herein called “inter-rib areas”. Each of the four big arches that cuts the vault in half defines a plane of symmetry (one half of the vault is the mirrored image of the other half). This symmetry holds in the three-dimensional space.

When the perspectival image in Saenredam’s painting is compared with a precise projection of the actual church (from the same viewpoint), it can be seen that the painter has extended the heights of the two large pointed arches on either side of the crossing [27]. Given his command of the techniques of perspective, this extension is probably not accidental. It probably arose from his sense that the vertical grandeur of the building could best be expressed by departing from optical accuracy.

In the case of planar structures three-dimensional symmetries translate, on the image plane, into slightly more complicated mathematical relationships. Those are known as *planar homologies* [15, 20, 25, 33, 37]. A homology is a plane-to-plane projective mapping defined by a line of fixed points called *axis*, a fixed point called *vertex* and a *characteristic* cross-ratio. It can be parametrized algebraically by a 3×3 matrix of the form:

$$\mathtt{H} = \mathtt{I} + \mu \frac{\mathbf{v}\mathbf{a}^\top}{\mathbf{v} \cdot \mathbf{a}} \tag{5}$$

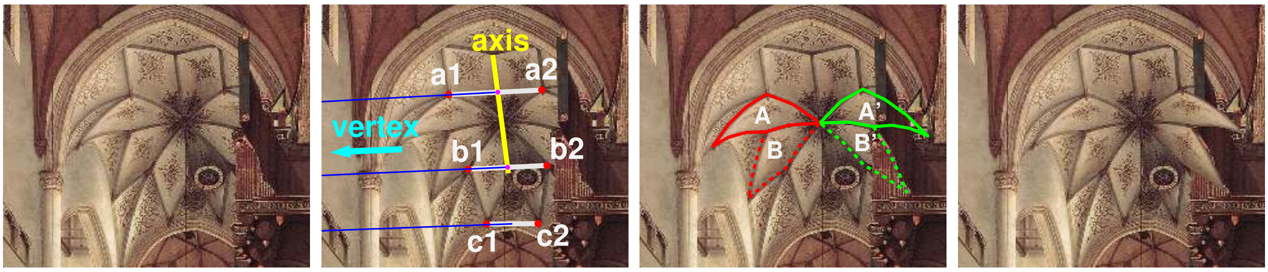
where $\mathbf{a}$ and $\mathbf{v}$ are 3-vectors representing the axis and the vertex (in homogeneous coordinates), respectively. μ is the scalar characteristic cross-ratio.

In the case of an imaged planar bilateral symmetry, corresponding points are related by a simpler transformation known as a *planar harmonic homology* (see fig. 13). This is a special case of a homology where the cross-ratio is known to be $\mu = -2$. Therefore:

$$\mathtt{H} = \mathtt{I} - 2 \frac{\mathbf{v}\mathbf{a}^\top}{\mathbf{v} \cdot \mathbf{a}} \tag{6}$$

It has the property that $\mathtt{H}^2 = \mathtt{I}$, where $\mathtt{I}$ is the identity matrix.

In fig. 14a one of the four arches is slightly different from the others in that its image is an almost perfect straight line. This line herein referred to as “axis” is the intersection of one of the planes of symmetry of the vault with the plane of the canvas and is plotted in yellow in fig. 14b. In the image plane the axis divides the vault into two sets of four inter-rib areas each. Corresponding inter-rib areas on either side of the vault are in a symmetry relationship in the three-dimensional space. If we approximate each of the eight inter-rib areas with a planar polygonal then the



a b c d

Figure 14: *Reconstructing occluded texture.* (**a**) *The vault of the Cathedral of St Bavo. Cropped from fig. 5.* (**b**) *The three-dimensional symmetry of the vault defines a planar harmonic homology between the inter-rib areas on the image plane. The axis of the homology is shown in yellow. Its vertex is outside the image, towards the left. The pairs of homologous points* ($\mathbf{a}_1, \mathbf{a}_2$), ($\mathbf{b}_1, \mathbf{b}_2$) *and* ($\mathbf{c}_1, \mathbf{c}_2$) *used to compute the vertex are shown in white.* (**c**) *Once the homology mapping has been computed, symmetrical patterns can be reconstructed in the image plane. For each pair of corresponding polygonals (A' is homologous to A and B' is homologous to B) the polygonal on the left (in red) has been manually selected and the one on the right (in green) has been automatically computed via the homology.* (**d**) *The occluded portions of the vault have been filled in by making use of the homology transformation. Pixel values in hidden areas (inside the two polygonals on the right, A and B') have been copied from the homologous visible pixels (in the corresponding polygonals on the left, A and B).*

geometrical relationship between corresponding inter-rib areas in the image plane can be approximated by a planar harmonic homology. This is only an approximation but we will show it to be sufficient for the purposes of this section.

The axis line in fig. 14b is actually the axis $\mathbf{a}$ of the homology we want to compute (*cf.* eq. 6). The vertex can be calculated easily by intersecting segments joining corresponding points on the two halves of the vault. In fig. 14b three pairs of corresponding points were selected on both sides of the vault. The vertex $\mathbf{v}$ is computed as the intersection of the lines defined by the pairs of corresponding points; in homogeneous coordinates $\mathbf{v} = (\mathbf{a}_1 \times \mathbf{a}_2) \times (\mathbf{b}_1 \times \mathbf{b}_2)$ or, equivalently $\mathbf{v} = (\mathbf{c}_1 \times \mathbf{c}_2) \times (\mathbf{b}_1 \times \mathbf{b}_2)$. The cross-ratio of the harmonic homology is known and therefore the matrix $\mathtt{H}$ in (6) is completely defined.

The computed harmonic homology $\mathtt{H}$ is now employed to map points on one side of the vault into the corresponding points on the other side. A point $\mathbf{p}$ on the left side of the vault is mapped into its homologous point $\mathbf{p}'$ via the homology $\mathtt{H}$ as follows:

$$\mathbf{p}' = \mathtt{H}\mathbf{p} \tag{7}$$

where $\mathbf{p}$ and $\mathbf{p}'$ are homogeneous representations (3-vectors) of image locations, i.e. for a generic point $\mathbf{q}$ we have $\mathbf{q} = (q_x, q_y, 1)^\top = \alpha(q_x, q_y, 1)^\top$.

In fig. 14c the contours of two inter-rib areas on the left side of the vault have been manually selected (shown in red) and, by applying the homology transformation, the corresponding polygonals on the right have been computed (shown in green). In fig. 14d the complete appearance of the vault has been recovered in the hidden areas by computing, for each pixel behind the organ its corresponding location on the visible areas and copying over its colour. Here all eight inter-rib areas of the vault are now visible.

Other possible texture-filling techniques include non-parametric texture synthesis [12] which are applicable only to front-on images and, therefore, may be applied to portions of paintings after an initial rectification step performed as described in section 3.1.

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); filling in occluded patterns. 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 [6].

a b c d e

Figure 15: *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. The reconstruction algorithm is described in [6]. (***b-e***) Different views of the reconstructed three-dimensional model of the chapel in the Florentine fresco.*

This section presents some of the three-dimensional models that have been reconstructed and explores the possible ambiguities in reconstruction that may arise[6]. Furthermore, three-dimensional reconstruction is used also as a tool to detect and magnify possible imperfections in the geometry of the painting.

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. 15a). 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 *entablurature*s 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

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

reach a consensus over those debated disputes. Views of the resulting three-dimensional model are shown in figs. 15b–e. The model geometry has been computed and stored in a VRML format which can be displayed by any of the many existing browsers[7].

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 section 3.4, here we have chosen not to do so, in order to highlight the areas which are not visible from the original viewing position.

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. 15b-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. 4.1 for both cases demonstrate that the square-ground-plan assumption yields rectangular coffers and the square-coffers assumption yields a rectangular ground plan. Views of the complete three-dimensional models corresponding to the two assumptions are shown in fig. 17.

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 2.3.

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.

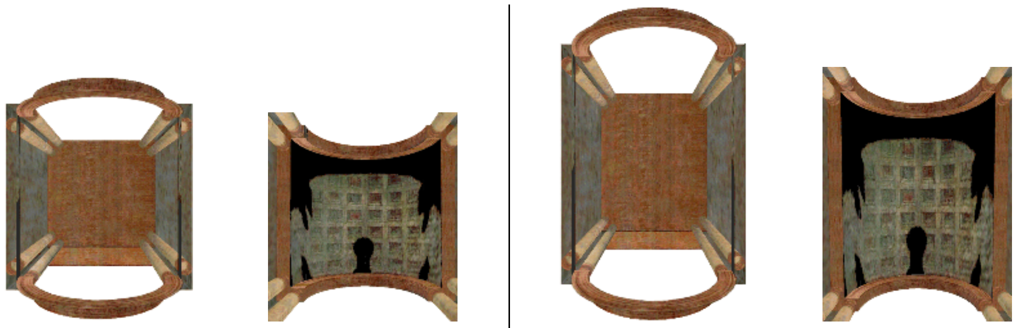
At this point new questions arise: Which architectonical structure wanted the artist to convey? If he had started by laying down a square base, why would he choose rectangular-shaped coffers? Was he aware of the depth ambiguity? Was it done on purpose?

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

---

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

Reconstruction assuming **square ground plan**		*Reconstruction assuming* **square vault coffers**	



floor:	*vault:*	*floor:*	*vault:*
square ground plan	rectangular coffers	rectangular ground plan	square coffers

Figure 16: *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.*

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 [9]. 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 [38], 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 [8] 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

Figure 17: *Two images of the models reconstructed from the* Trinity *fresco. (***a***) The three-dimensional model originated from the assumption of square ground plan. (***b***) The three-dimensional model corresponding to the assumption of square coffers. Both models are geometrically consistent in that both can be correctly projected onto the same flat image, i.e. the original fresco, but the latter is deeper.*

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.

4.2 The virtual *Flagellation*

Piero's *Flagellation* (in fig. 7a) conforms faithfully to the rules of linear perspective. Therefore, the depicted scene can correctly be reconstructed in the three-dimensional space.

Figure 18 show views of the reconstructed VRML model taken from different viewpoints. In this case only one assumption has been made, i.e. that the black and white pattern on the floor is square. From the recovered 3D geometry this assumption appears to be consistent with that of a square and regular pattern in the ceiling of the architectural structure supported by the columns (in fig. 18c).

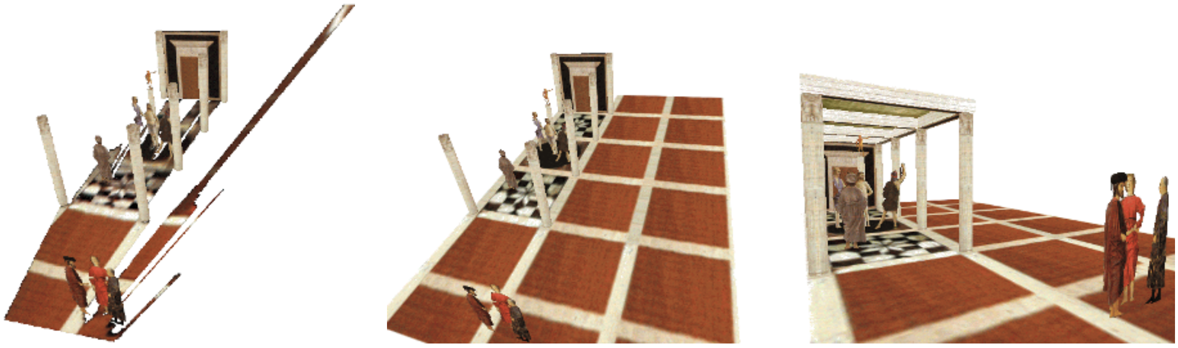
Captivating visual experiences arise from the capability of turning flat paintings into three-dimensional objects and exploring them in a "virtual reality" fashion. One question arises: How can we ensure that the reconstructed model is correct? Ways to assess the accuracy of the constructed three-dimensional model and its consistency with the original painting will be presented in section 5.

4.3 The virtual *St Jerome*

St Jerome in His Study (see fig. 19a) is an oil painting by the Dutch artist H. Steenwick (1580–1649), who was one of the pioneers of perspectival interiors in Dutch painting.

Linear perspective was generally adopted later in northern Europe than in Italy, but it was in Holland, where elaborate depictions of buildings and townscapes in their own right became a major genre for painters in the seventeenth century, that the potential of Brunelleschi's invention for the depiction of actual (or apparently real) views was fully realised. Saenredam stands as one of the supreme heirs of Steenwick and early theorists of perspective in Holland, such as Hans Vredeman de Vries (*cf.* [22] pp. 109-111).

The accuracy of the perspective in Steenwick's *St Jerome* and the amazing management of light shade as it traverses the spaces make of this painting a very significant, early example of Dutch painting of domestic and ecclesiastical interiors [27] , combining in this case both a room and a distant vista into a church. The beautifully characterised



a b c

Figure 18: *Three-dimensional reconstruction of Piero's* Flagellation. *(***a***) A view of the reconstructed three-dimensional model with the roof removed to show the relative positions of people and architectural elements in the scene. The occluded regions on the floor have been left blank and the people are represented as flat silhouettes. The columns have been approximated by cylinders. Details about the reconstruction algorithm are in [6]. (***b***) The occluded areas of the tiled floor have been filled in by taking advantage of the symmetry of the floor pattern. (***c***) The partially visible ceiling has also been reconstructed.*

passage of the light from the windows on the left, casting shadows across the tiled floor, gives Steenwick's imagined interior an extraordinary sense of veracity.

Given its strong geometrical component (numerous parallel lines and planar surfaces can be observed) the painting proves an ideal input for our reconstruction techniques.

Views of the models obtained by applying our computer-based reconstruction algorithms to the original painting are illustrated in figs. 19b–f. This three-dimensional model looks particularly realistic when displayed in a virtual cave [8]. The virtual cave offers the user the possibility to physically enter the painted scene, fly around it and "virtually" become part of the painting itself.

The following section shows how reconstructing a painting in three dimensions offers the possibility to detect and investigate inconsistencies which are hard to notice through an analysis of the flat original image only.

Detecting geometric inconsistencies.

Figure 20a shows a front-on view of the reconstructed large window on the left hand side of the painting. The pattern of the glass window has been completed in its occluded areas by making use of symmetry relationships as in section 3.4.

Notice that while parallelism and angles have been recovered correctly (in fig. 20a the window pattern looks correctly rectified) an unexpected, asymmetric curvature of the top arch can be detected. The right side of the arch appears to be thicker than the left side (the highlighted area in fig. 20b). This inconsistency is made evident by our reconstruction process and is less noticeable in views taken from locations closer to the original viewpoint (*cf.* fig. 20c–f).

This geometrical imperfection is probably due to the fact that the artist has painted a somehow complicated curve at a grazing angle by eye and without undertaking a precise projection, which would have promised a degree of effort disproportionate to the visual benefits (*cf.* the original painting in fig. 19a). A large uncertainty, and therefore risks of inaccuracies, characterizes such a situation. This is analysed, below, by making use of rigorous statistical tools.

The uncertainty in the location of points can be modelled as a *Gaussian* random variable and described by a 2×2 covariance matrix. The covariance in the location of a point feature can be visualised as an ellipse [6].

Here we make the assumption that during the painting process the probability of the artist making a mistake is isotropic and uniformly distributed across the plane of the painting, i.e. there is the same likelihood of making a mistake in any direction and in any point of the canvas.

[8] This three-dimensional model and others have been installed at the virtual cave (the vr-cube) that is being developped at Stockholm's KTH. For further details go to `www.pdc.kth.se/projects/vr-cube/`.

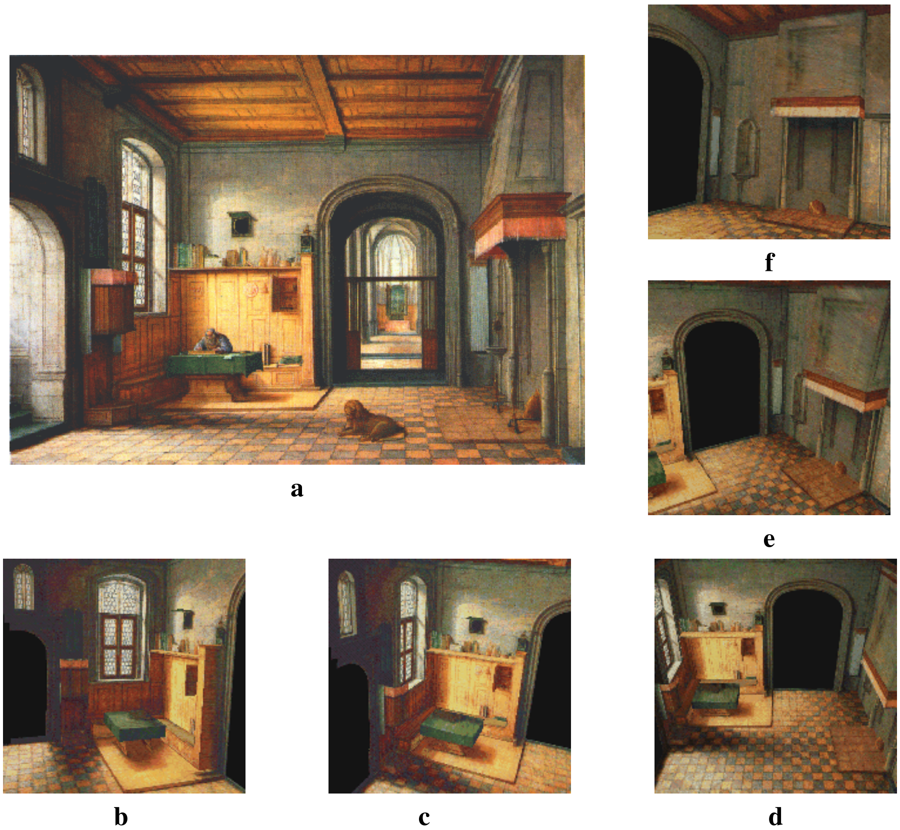


Figure 19: *Three-dimensional reconstruction of* St Jerome in his Study. *(***a***)* St Jerome in his Study *(1630), by Hendrick V. Steenwick (1580–1649), private collection (Joseph R. Ritman Collection), Amsterdam, The Netherlands. (***b-f***) Different views of a "virtual fly-through" inside the computed three-dimensional model.*

In figure 21a the distribution of the uncertainty on the plane of the painting is visualised by superimposing a regular grid of circles on the original painting. In particular, we are only interested in the regions of the large window on the left.

In general, a uniform and isotropic distribution of uncertainties on the canvas plane translates into non-uniform and non-isotropic uncertainties on the reconstructed three-dimensional scene. Figure 21b shows a front-on view of the window, computed with the usual method of rectifying the image by applying a homography transformation to the original painting (*cf.* fig. 10). The circles in fig. 21a are mapped by the reconstruction process into ellipses of varying size and orientation. This demonstrates that the uncertainty in the reconstructed model is not isotropic (the ellipses do not have circular symmetry) nor is uniform (the ellipses are denser towards the left-hand side of fig. 21b). The increase in the ellipses size going from left to right in fig. 21b accounts for the reduced accuracy of the right side of the window arch.

The idea of investigating geometric imperfections by generating new views of portions of a painting was already present in what is considered to be the very first treatise on perspective, i.e. "*Della Pictura*" by Leon Battista Alberti [1]. In his treatise Alberti suggested looking at paintings in a mirror to expose its weaknesses [9].

The kind of inconsistencies that arise in such details in Steenwick's carefully-constructed painting are related to what psychologists of perception call the "constancy" factor, that is to say our tendency to judge seen forms as less foreshortened and diminished in size than the geometry of perspective demands. Thus we tend to "see" the top of a cup viewed obliquely as less elliptical and more round than a purely optical projection would produce. Left to paint

[9] "*It is marvelous how every weakness in a painting is so manifestly deformed in the mirror.*"

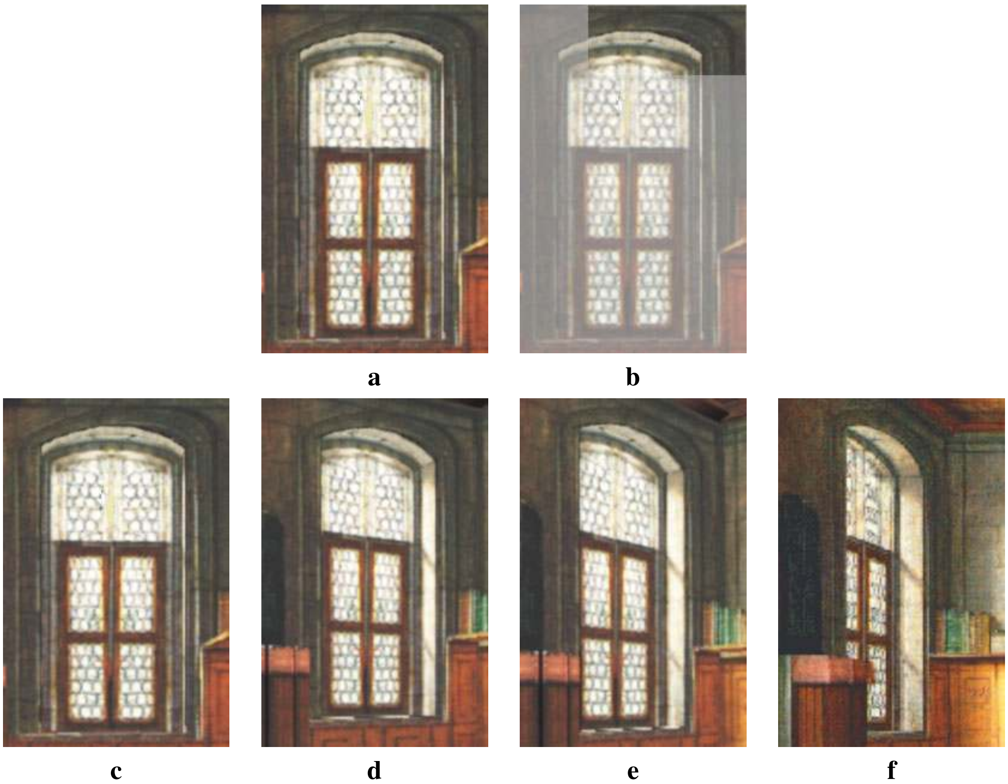


Figure 20: *Detecting geometric inconsistencies in the painting* St Jerome in his Study *by comparing two views of a reconstructed window. (***a***) a fronto-parallel view of the window on the left wall of the reconstructed model. (***b***) The arch appears asymmetric. The right side of the arch (highlighted) is thicker than the left side. (***c-f***) The sequence of four images show how the asymmetry of the window's arch is less noticeable when observed from an increasingly grazing angle, closer to the original eye point. Images* **(c-e)** *are captured from the reconstructed 3D model and* **(f)** *is cropped from the original painting.*

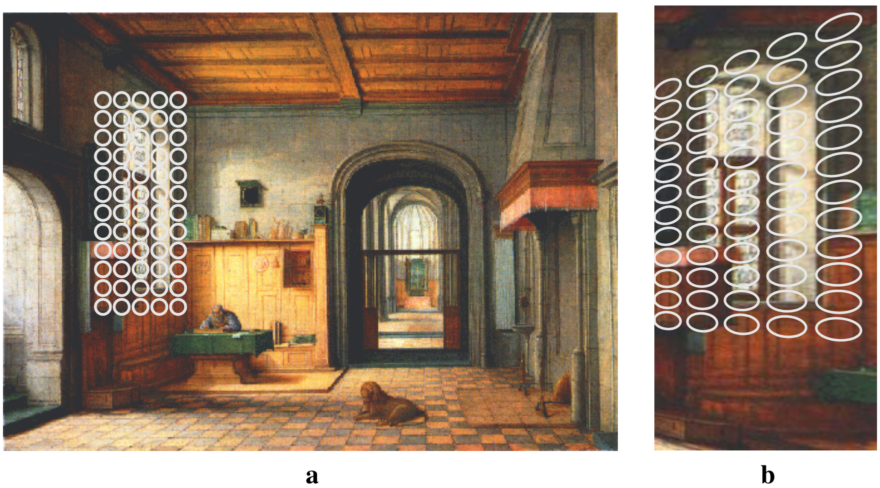


Figure 21: *Analysis of the accuracy of the geometry of a painting. (***a***) The distribution of the uncertainty in a painting is assumed to be uniform and isotropic. This is represented by the overlaid uniform grid of circles. (***b***) The rectified image of the window. The circles in (a) are mapped into ellipses, thus indicating the loss of isotropy and uniformity with respect to the original uncertainty distribution. Larger uncertainties (larger ellipses) characterize the right-hand side of the rectified window. See text for details.*

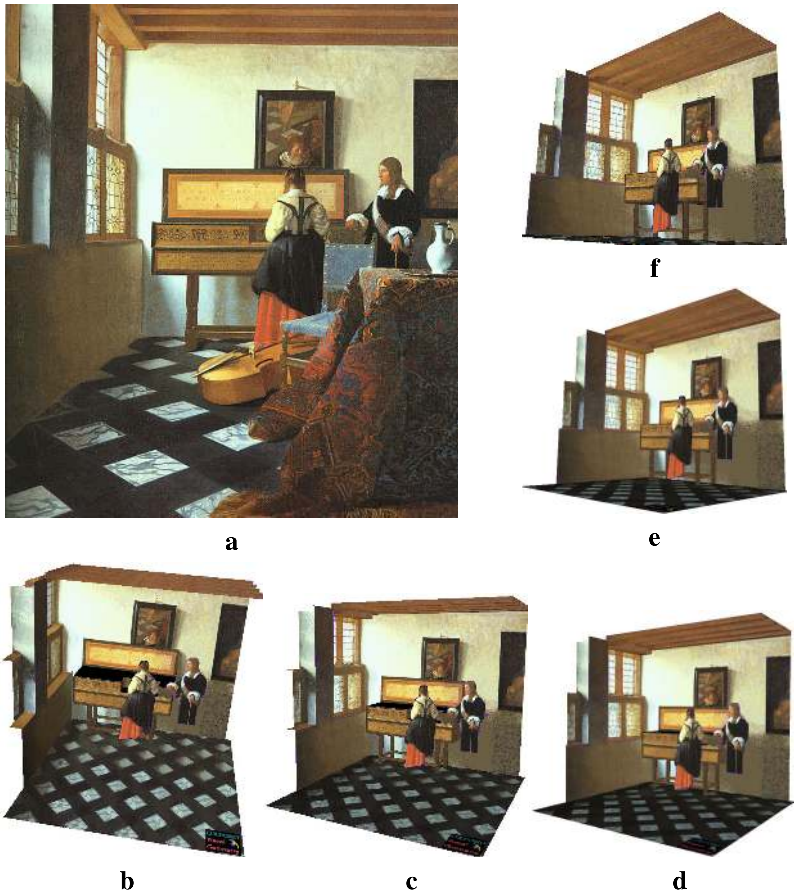


Figure 22: *Three-dimensional reconstruction of* The Music Lesson. *(***a***)* A lady at the virginals with a gentleman (the music lesson) *(1662-65),* 73.3 × 64.5 *cm, by Johannes Vermeer (1632-1675). (***b-f***) Snapshots of a virtual fly-through inside the reconstructed painting to show different views of the room.*

something "by eye", as Steenwick was done with the window, there is a tendency to make the forms "look right" according to perspectual rather than geometrical criteria [39].

4.4 The virtual *Music Lesson*

The Music Lesson by Jan Vermeer (fig. 22a) is another example of the depiction of compelling interiors in Dutch paintings. A perfect geometric construction of the depicted scene (probably a real room) is coupled together with an unparalleled skill in the characterisation of light. Such is Vermeer's concern with the behaviour of light, shade and colour as it effects our eye that he seems to be painting with light as much as with geometry, in a way that none of his contemporaries achieved [29].

The three-dimensional model constructed by applying the single-view techniques is shown in fig. 22. Here we have restricted the reconstruction model to mostly piecewise planar, and therefore complex shapes such as the cello or the decorated table cloth have not been considered in the final model. For such free-form shapes, techniques based on a "three-dimensional repainting" paradigm are more indicated [21, 28, 38]. On the other hand those free-form repainting methodologies are as prone to personal interpretation as the original painting process itself and are not appropriate when geometric accuracy is a requirement.

The results are consistent with those achieved by Steadman [35], who has used traditional modes of analysis to construct an actual model of the room, in order to test the match between Vermeer's painted space and an actual

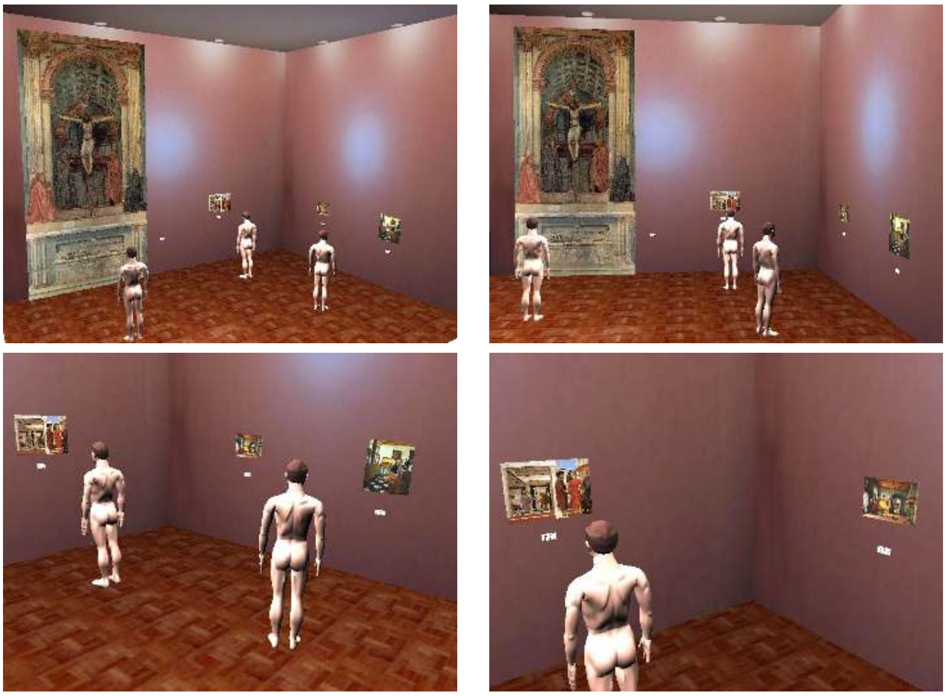


Figure 23: *Snapshots of our 3D interactive virtual museum: an XML-based tool which allows the user to visualise flat paintings in a virtual museum environment and seamlessly "dive" into the paintings third dimension. The four paintings are ordered chronologically from left to right. The Italian* Trinity *and* Flagellation *are hung on the left wall. The Dutch* St Jerome *and* The Music Lesson *are hung on the right wall. The paintings are in scale and at a suitable height for the eye point of a notional spectator. Note that the* Trinity *fresco is far larger than the other three paintings. Avatars have been inserted to better convey the sense of scale.*

room with realistically scaled objects and real light sources. It would be possible to combine Steadman's empirically constructed cello, and the table with its covering and contents to construct an entirely consistent virtual model of the room which Vermeer seems to have used in at least eight of his paintings. The analysis presented here strongly supports the view that Vermeer set up the rear side of the room as a camera obscura to lay down the basic illusion in his paintings.

5 An Interactive Virtual Museum

This section presents an interactive visualization tool that enables the user to visualise paintings in a three-dimensional environment and, more importantly, interact with them by "diving" into the corresponding three-dimensional scenes, thus allowing a fuller visual experience of the painted space and providing an insight into the artists' modes of visualisation and presentation.

The four paintings and frescoes analysed in the previous section have been virtually hung on the walls of a three-dimensional virtual museum in accordance with their relative scales and at a suitable height for they eye-point of a notional spectator (in fig. 23). The difference between our application and other (possibly web-based) virtual museums is the possibility for the user to experience the works of art in a completely new and exciting way, by "diving" into the painting's third dimension. Our virtual museum allows the user to move freely within the museum room and, also, from the museum room into the three-dimensional scenes reconstructed from the paintings.

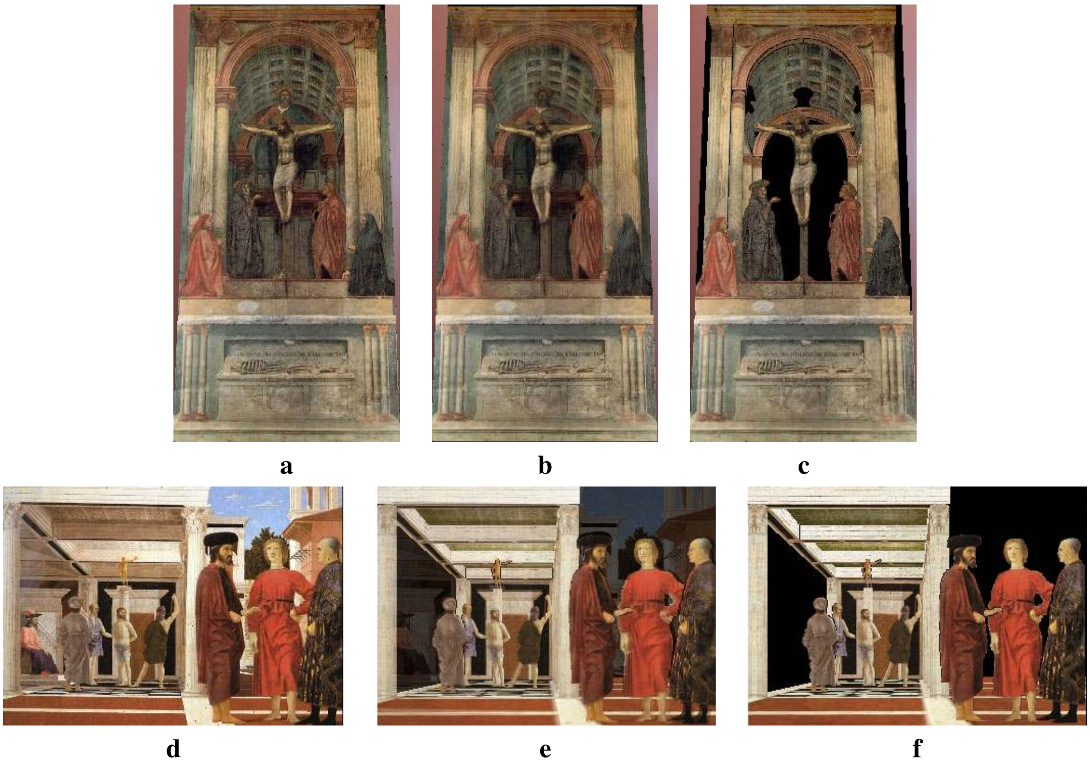


Figure 24: *Seamless transitions between the virtual museum and the reconstructed 3D paintings. (***a***), (***b***) and (***c***) The fading transition, from a view of the* Trinity *fresco (a) to a view of its reconstructed 3D model from the correct vantage location (c). Image (c) is the* reprojection image *for the* Trinity*, see text. (***d***), (***e***) and (***f***) The fading transition, from a view of the* Flagellation *painting (d) to a view of its reconstructed 3D model (f). Image (f) is the* reprojection image *for the* Flagellation*, see text. These transitions are smooth and seamless because they are performed from each painting's vantage point.*

5.1 "Diving" into the paintings' third dimension

In order to create compelling visual effects the transition from the museum room to the three-dimensional painting must be smooth and seamless. This can be achieved only if the following two constraints are matched: (i) the geometry of the reconstructed painting is accurately recovered, and (ii) the best viewing position is estimated correctly.

Each perspective painting is characterized by a location from which the observer experiences the most complete three-dimensional illusion, namely the *vantage point*. Any other location would not work as well for the observer, though, in practice, the willingness of our perceptual system to collaborate with the painted illusion means that the effect of the perspective remains robust for a greater range of viewing positions than is theoretically desirable.

In computer vision, computing the vantage location is equivalent to recovering the camera pose. The single-view reconstruction techniques enable us to compute both the three-dimensional geometry and the vantage point of each painting quite accurately.

In our virtual museum the desired quality of transition is achieved by forcing the observer to enter the paintings three-dimensional scenes only from the correct vantage locations. Furthermore, in order to enhance the visual realism of the transitory phase a simple fading between the painting and the three-dimensional scene behind the canvas is implemented at each entry point (*cf.* fig. 24).

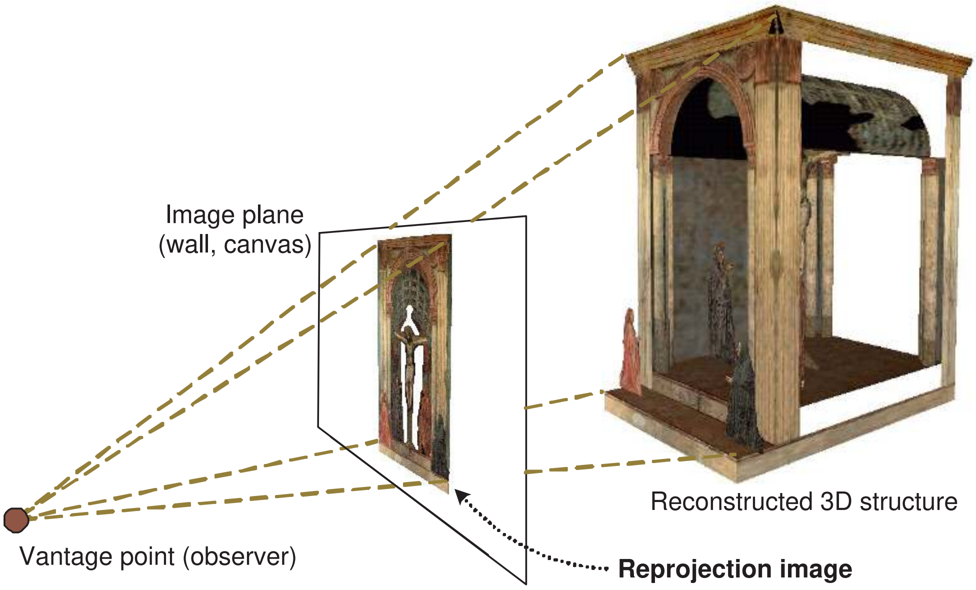


Figure 25: *The reprojection image is defined as the image created by projecting the computed three-dimensional reconstruction onto the plane of the painting.*

5.2 Accuracy of three-dimensional reconstruction

The above section suggests the use of our three-dimensional visualization tool as a way of assessing the accuracy of the reconstructed scene geometry and the correctness of the retrieved vantage position. In fact, the perfect transition from the museum room to the painted scene is achieved only when both the reconstructed 3D painting and its vantage location have been computed correctly. Otherwise, inaccuracies in either the recovered structure or the vantage point would show during the transition as misalignment and ghostly effects. Ghostly effects are visual instances of what in the computer vision field is commonly known as the *reprojection error* [15].

Reprojection error is due to misalignment between the original image, i.e. the painting, and the image which originates from projecting the computed three-dimensional scene onto the plane of the painting or fresco, namely the *reprojection image* (*cf.* fig. 25).

This section defines a measure of misalignment as a function of the original painting and the reprojection image. The misalignment is used as a measure of accuracy of a painting's three-dimensional reconstruction.

Given a painting, the reconstructed model and the computed vantage position it is easy to obtain the reprojection image by projecting the three-dimensional model onto the plane of the painting (*cf.* fig. 25).

How can we measure the misalignment between a painting and the corresponding reprojection image (see figs 24a,c and figs 24d,f)? One may think that simple image differencing would do, but it turns out that this is not a good solution. In fact, the difference between the two images would be large not only where the images are misaligned but also in those areas where there is a large difference in lighting. Since our reconstruction techniques focus on achieving a correct geometrical reconstruction rather than faithfully reproducing the correct lighting conditions, it is preferable to design a misalignment measure that took into account only variations in object positions, rather than their appearance (i.e. colour and shading). In order to achieve this goal we have decided to use a misalignment measure based on distances between object contours. Specifically, we make use of the Chamfer distance between two sets of contours [3].

Figure 26a shows the results of running a Canny edge detector [5] on the image of Masaccio's *Trinity* (fig. 24a). The black segments are the loci of the image where a sharp change in the image intensity is occurring. These edges are located at object boundaries and texture markings. Instead, fig. 26b shows the edges extracted from the reprojection image (fig. 24c).

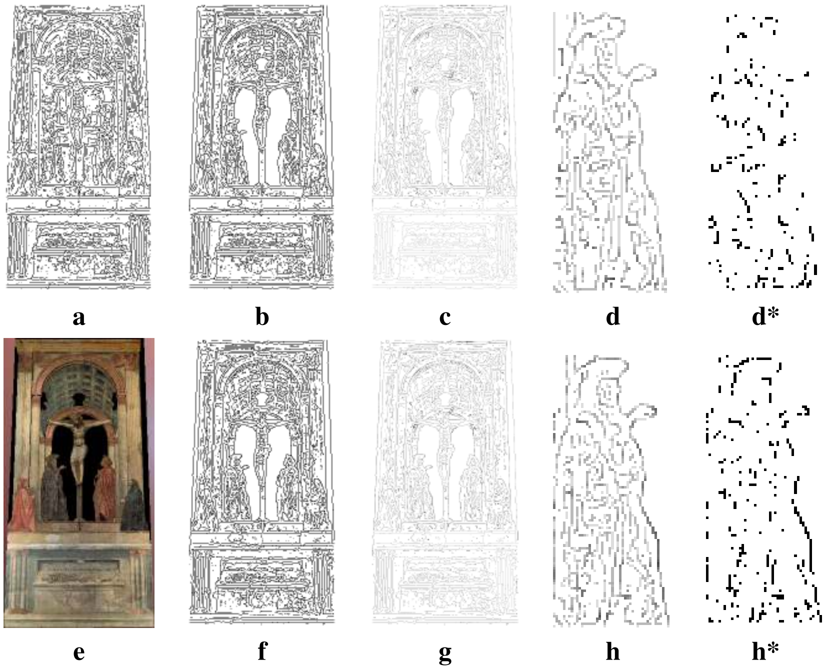


Figure 26: *Edge distance to compute misalignment error. (***a***) Canny edges detected on the image of the* Trinity. *(***b***) Canny edges detected on the reprojection image of the* Trinity. *(***c***) Chamfer distance between image (b) and (a). (***d***) Detail of (c). (***d****) As in (d) where only the strongest edges (greatest mismatches) have been maintained (in black) for ease of visualization. (***e***) Reprojected image related to the model where some figures have been misplaced on purpose: The cross has been pushed back, and the figures of the Virgin and St John the Baptist have been translated towards each other. (***f***) Canny edges detected on the reprojection image of the* Trinity *after the introduced 3D misplacement. (***g***) Chamfer distance between the edges in (e) and those in (a). (***h***) Detail of (g), to be compared with (d). (***h****) As in (h) where only the strongest edges (greatest mismatches) have been maintained (in black, see, for instance, the rightmost edge of the Virgin's dress). The edges in (h) are, in average, darker than the ones in (d) indicating larger misalignment. This is also observed by comparing (h*) with (d*): more mismatches (more black points) are present in (h*).*

The Chamfer image between figs 26a and 26b is shown in fig. 26c and it represents the misalignment between the painting and the corresponding reprojection image. The edges in the Chamfer image are intensity coded: the intensity values are proportional to the distance of each feature point in fig. 26b from the closest feature point in fig. 26a. The lighter the intensity, the smaller the distance. Therefore, in fig. 26c darker values indicate worse alignment (see fig. 26d for a detail). For a perfect alignment between the original painting and its reprojection image the Chamfer image would be completely white.

The misalignment between the image of a painting and its reprojected image can be computed, numerically, simply as the average of the values in the Chamfer image. This is the standard definition of Chamfer distance.

For example, for the figures 24a and 24b the misalignment has been computed to be

$$Mis = 0.72 \;\; pix \tag{8}$$

Larger values of (8) mean a larger misalignment between the original and reprojected images and is an indication of inaccurate three-dimensional reconstruction (inaccurate scene structure, or incorrect computation of the vantage point). Moreover, the Chamfer image shows darker edges in those places where a larger misalignment occurs. This effect may be used to pinpoint where mistakes are occurring and, therefore, interactively refine the reconstructed model.

For a better understanding of the defined misalignment measure we run a test in the following steps:

1. given a painting's three-dimensional model we randomly select a few objects and translate them to new locations;

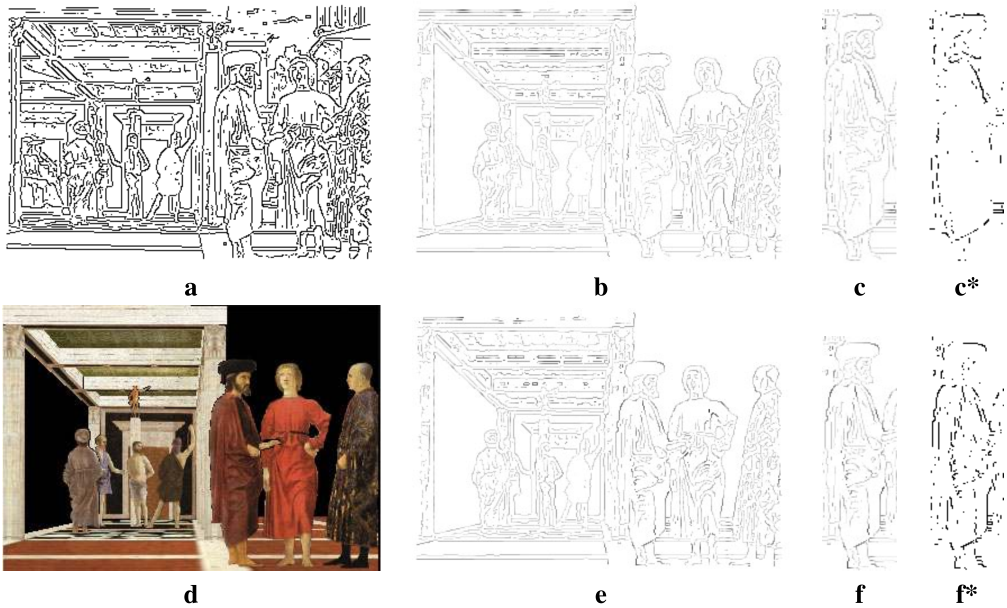


Figure 27: *Edge distance to compute misalignment error. (***a***) Canny edges detected on the image of the* Flagellation*. (***b***) Chamfer image between the edges of the reprojected image and the ones in (a). (***c***) Detail of (b). (***c****) As in (c) where only the strongest edges (greatest mismatches) have been maintained (in black) for ease of visualization. (***d***) Reprojected image related to the model where some figures have been misplaced on purpose: the figure of Christ has been moved towards the right and two of the foreground figures towards the left. (***e***) Chamfer image between the edges extracted from (c) and those in (a). (***f***) Detail of (e). (***f****) As in (f) where only the strongest edges (greatest mismatches) have been maintained (in black). The edges in (f) are, in average, darker than the ones in (c) indicating larger misalignment. This is also observed by comparing (f*) with (c*): more mismatches (more black points) are present in (f*).*

2. we construct the new reprojection image (from the same vantage viewpoint);
3. finally we construct the Chamfer image and measure the misalignment.

For instance, in the model of the *Trinity* we may push the cross backwards and the figures of the Virgin and St John the Baptist closer to each other. The corresponding reprojection image is shown in fig. 26e. This is slightly different from fig. 24c because of the introduced misalignment. The extracted Canny edges are shown in fig. 26f and the Chamfer image between fig. 26f and the edges of the original painting (fig. 26a) is shown in fig. 26g. In this case the misalignment value has been computed to be

$$Mis' = 0.91 \ \ pix$$

The fact that $Mis' > Mis$ indicates a worse alignment between original and reprojected images, as expected.

Notice that the defined misalignment measure is non-convex, which means that it may not behave correctly in situations where large displacements occur. This problem may be mitigated by first running an edge matching preprocessing step [32, 34] and then measuring the average distance between corresponding edges.

A similar experiment was run on the *Flagellation* and the results are shown in fig. 27. In this case the misalignment between the original set of edges and those extracted from the reprojected image (fig. 27b) was computed to be:

$$Mis = 1.06 \ \ pix$$

By purposely translating the figure of Christ towards the right and two of the foreground figures towards the left we obtained the reprojected image shown in fig. 27d and the related Chamfer image in fig. 27e. In this case the misalignment between painting and reprojected image has been computed as:

$$Mis' = 1.23 \ \ pix$$

Once again $Mis' > Mis$ as expected; thus confirming the validity of this function as a measure of the accuracy of the three-dimensional reconstruction. A similar effect is obtained by moving the viewing position away from the vantage location.

6 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*? [10] What does the pattern on the floor in Domenico Veneziano's *St Lucy Altarpiece* (*cf.* fig 11) look like? What is the shape of the dome in Raphael's *School of Athens* (*cf.* fig 12)?

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 *Empire of the Eye* project under design for the National Gallery in Washington[11]. *Empire of the Eye* intends to build upon and extend *Masters of Illusions* [19]. 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 [26].

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 [22] it should be possible to take some of the artists who adopted a consciously experimental approach to light and colour, such as the

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

[11] `www.nga.gov`

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.

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[12]. However, as a large amount of psychophysics research shows [2, 4], our brain is capable of neglecting conflicting cues, and slightly inaccurate perspective or shading may 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] Leon Battista Alberti. *De Pictura*. 1435. Reproduced by Laterza (1980).

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

[3] G. Borgefors. Hierarchical chamfer matching: A parametric edge matching algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(6):849–865, 1988.

[4] 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.

[5] J. F. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.

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

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

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

[12] as seen in fig.3 images of parallel lines may not intersect in the exact same point

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

[10] 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.

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

[12] A. Efros and T. Leung. Texture syntesis by non-parametric sampling. In *Proc. 7th International Conference on Computer Vision, Kerkyra, Greece*, pages 1039–1046, September 1999.

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

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

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

[16] D. Hockney. *That's the way I see it*. Nicos Stangos, Thames and Hudson, London, 1993.

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

[18] 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.

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

[20] T. Kanade. Recovery of three-dimensional shape of an object from a single view. *Artificial Intelligence*, 17(1–3):409–460, 1981.

[21] S. B. Kang. Depth painting for image-based rendering applications. Technical report, CRL, Compaq Cambridge Research Lab, 1998.

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

[23] 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.

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

[25] D. P. Mukherjee, A. Zisserman, and J. M. Brady. Shape from symmetry—detecting and exploiting symmetry in affine images. In *Philosophical Transactions of the Royal Society of London, SERIES A*, volume 351, pages 77–106, 1995.

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

[27] National Gallery of Scotland. *Dutch Church Painters*. The National Galleries of Scotland, Edinburgh, 1984.

[28] B. M. Oh, M. Chen, J. Dorsey, and F. Durand. Image-based modeling and photo editing. In *Proc. ACM SIGGRAPH*, Los Angeles, CA, USA, 2001.

[29] S. Peccatori and S. Zuffi. *ArtBook Vermeer: The master of light and composition - his life in paintings*. Dorling Kindersley, 1999.

[30] R. Ruurs. Drawing from nature. *La Prospettiva Rinascimentale*, 1980.

[31] F. Schaffalitzky and A. Zisserman. Planar grouping for automatic detection of vanishing lines and points. *Image and Vision Computing*, 18:647–658, 2000.

[32] C. Schmid and A. Zisserman. The geometry and matching of curves in multiple views. In *Proc. European Conference on Computer Vision*, pages 394–409. Springer-Verlag, June 1998.

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

[34] Y. Shan and Z. Zhang. New measurements and corner-guidance for curve matching with probabilistic relaxation. *International Journal of Computer Vision*, 46(2):157–171, 2002.

[35] P. Steadman. *Vermeer's Camera: Uncovering the Truth Behind the Masterpieces*. Oxford University Press, Oxford, 2001. ISBN: 0-19-215967-4.

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

[37] L. van Gool, M. Proesmans, and A. Zisserman. Planar homologies as a basis for grouping and recognition. *Image and Vision Computing*, 16:21–26, January 1998.

[38] 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.

[39] D. N. Zorin. *Correction of Perceptual Distortions in Pictures*. PhD thesis, California Institue of Technology, Pasadena, CA, USA, 1995.