Accidental Pinhole and Pinspeck Cameras

Revealing the Scene Outside the Picture

Antonio Torralba · William T. Freeman

Abstract We identify and study two types of "accidental" images that can be formed in scenes. The first is an accidental pinhole camera image. The second class of accidental images are "inverse" pinhole camera images, formed by subtracting an image with a small occluder present from a reference image without the occluder. Both types of accidental cameras happen in a variety of different situations. For example, an indoor scene illuminated by natural light, a street with a person walking under the shadow of a building, etc. The images produced by accidental cameras are often mistaken for shadows or interreflections. However, accidental images can reveal information about the scene outside the image, the lighting conditions, or the aperture by which light enters the scene.

Keywords Accidental cameras · Pinhole · Anti pinhole

1 Introduction

There are many ways in which pictures are formed around us. The most efficient mechanisms are to use lenses or narrow apertures to focus light into a picture of what is in front. A set of occluders (to form a pinhole camera) or a mirror surface (to capture only a subset of the reflected rays) let us see an image as we view a surface. Researchers in computer vision have explored numerous ways to form images, including novel lenses, mirrors, coded apertures, and light sources (e.g. Adelson and Wang 1992; Baker and Nayar 1999; Levin

A. Torralba (✉) · W. T. Freeman
Computer Science and Artificial Intelligence Laboratory,
Massachusetts Institute of Technology, Cambridge, MA 02139, USA
e-mail: torralba@csail.mit.edu

W. T. Freeman
e-mail: billf@mit.edu

et al. 2007; Nayar et al. 2006). The novel cameras are, by necessity, carefully designed to control the light transport such that images can be viewed from the data recorded by the sensors. For those cases, an image is formed by intentionally building a particular arrangement of surfaces that will result in a camera. However, similar arrangements appear naturally by accidental arrangements of surfaces in many places. Often the observer is not aware of the faint images produced by those accidental cameras.

Figure 1 shows a picture of a hotel room somewhere in Spain. There would be nothing special in this picture if it wasn't for the pattern of darkness and light on the wall. At first, one could mis-interpret some of the dark patterns on the wall of the bedroom as shadows. But after close inspection, it is hard to understand which objects could be casting those shadows on the wall. Understanding the origin of those shadows requires looking at the full environment surrounding that wall. Figure 2a shows a montage of the full scene. All the light inside the room enters via an open window facing the wall. Outside the room there is a patio getting direct sunlight. As there are no objects blocking the window and producing those shadows we will have to look for a different explanation for the patterns appearing on the wall. What is happening here is that the window of the room is acting as a pinhole and the entire room has become an *accidental pinhole camera* projecting an image onto the wall. As the window is large, the projected image is a blurry picture of the outside. One way to confirm our hypothesis and to reveal the origin of light patterns that appear in the room is to block the window to only allow light to enter via a narrow aperture, thus transforming the room into a camera obscura. After blocking the window, the projected image appears sharp as shown in Fig. 2c. Now we can see that the light patterns shown on Fig. 1 were not shadows but a very blurry upside-down image of the scene outside the room (Fig. 2e).

Fig. 1 What are the dark regions on the *white wall*? Are they shadows? See Fig. 2 to get the answer

Perceiving as images the light projected by a pinhole into a wall with an arbitrary geometry might not be easy, especially when the image is created by an accidental camera. This, together with blurring from the large window aperture,

leads to most such accidental images being interpreted as shadows. In this paper, we point out that in scenes, *accidental* images can form, and can be revealed within still images or extracted from a video sequence using simple processing, corresponding to accidental pinhole and "inverse" pinhole camera images, respectively. These images are typically of poorer quality than images formed by intentional cameras, but they are present in many scenes illuminated by indirect light and often occur without us noticing them.

Accidental cameras can have applications in image forensics as they can be used to reveal other parts of the scene not directly shown in a picture or video. Accidental images can be used to better understand the patterns of light seen on a normal scene that many times are wrongly identified as shadows. In the literature there are examples of accidental cameras being used to extract information not directly available in the original picture. For instance, the scene might also contain reflective surfaces (e.g. the faucet or a mirror) which might reveal a distorted image of what is outside of the picture frame. In Nishino and Nayar (2006) the authors show an example of *accidental mirrors*. They show how to extract an image of what is on the other side of the camera by analyzing the reflected image on the eye of the people present in

Fig. 2 An accidental pinhole camera: light enters a room via an open window. The window restricts the light rays that can enter the room, just as a pinhole camera does, creating a faint picture on the wall of the scene outside the room. (**a**) Montage of the full scene of the hotel room and the patio outside, (**b**) picture of the wall when the window is full open, (**c**) picture of the wall when the window is turned into a tiny pinhole. (**d**) Upside-down picture, (**e**) true view outside the window

Fig. 3 Relaxing the pinhole camera design. (**a**) Pinhole camera from a class project (the small thumbnail shows a picture taken with this camera). (**b**) Relaxing the design of the pinhole camera by removing the walls of the camera. (**c**) Turning the room into a camera obscura using whatever objects were around to reduce the opening. (**d**) Accidental creation of a pinhole. The pinhole is formed by the right arm against the body, an upside-down, faint and blurry picture of the window can be seen projected on the wall

the picture. A Bayesian analysis of diffuse reflections over many different times has been used for imaging in astronomy applications (Hasinoff et al. 2011).

In this paper we identify and study two types of accidental cameras (pinholes and antipinholes) that can be formed in scenes, extending the work described in Torralba and Freeman (2012). In Sect. 2 we review the principles behind the pinhole camera. We also describe situations in which accidental pinhole cameras arise and how the accidental images can be extracted from pictures. In Sect. 3 we discuss the anti-pinhole cameras and we show how shadows can be used as accidental anti-pinhole cameras revealing the scene outside the picture. In Sect. 4 we discuss applications and show examples of accidental cameras.

2 Accidental Pinhole Cameras

The goal of this section is to illustrate a number of situations in which accidental pinhole cameras are formed and to educate the eye of the reader to see the accidental images that one might encounter in daily scenes around us. We show how we can use Retinex (Land and McCann 1971) to extract the accidental images formed by accidental pinhole cameras.

2.1 Pinhole Camera

In order to build a good pinhole camera we need to take care of several details. Figure 3a shows a pinhole camera built for a class exercise. In this box there are two openings: one

large opening (clearly visible in the picture) where we can insert a digital camera and a small opening near the center that will be the one letting light inside the box. The digital camera will be used to take a long exposure picture of the image projected on the white paper. Light will enter via a small hole. The smaller the hole, the sharper the picture will be. The inside of the camera has to be black to avoid inter-reflections. The distance between the hole and the back of the box (focal length) and the size of the white paper will determine the angle of view of the camera. If the box is very deep, then the picture will correspond to only a narrow angle.

It is important to follow all those procedures in order to get good quality pictures. However, if one is willing to lose image quality, it is possible to significantly relax the design constraints and still get reasonable images. This is illustrated in Fig. 3. In Fig. 3b the pinhole camera has been replaced by two pieces of paper, one paper is white and it will be used to form an image and the other one has a hole in the middle. Now light arrives to the image plane from multiple directions as there is no box to block all the light rays that do not come from the pinhole. However, still an image gets formed and has enough contrast to be visible by the naked eye. Despite the low quality of the image, this setting creates a compelling effect as one can stand nearby and see the image projected. Figure 3c shows how the room is turned into a camera obscura without taking too much care on how the window is blocked to produce a small opening. In this case the window is partially closed and blocked with a pillow and some cushions. Despite that several openings are still present, a picture of the buildings outside the room gets projected on the wall. In Fig. 3d we see a more extreme situation in which now the pieces of paper have been replaced by a more *accidental* set of surfaces. In this case, a person stands in front of a wall. A small opening between the arms and body creates a pinhole and projects a faint image on the wall. The pinhole is not completely circular, but still creates an image.

The goal of these visual experiments is to help the viewer to get familiar with the notion that pinhole cameras can be substantially simplified and still produce reasonable images. Therefore, one can expect that these more relaxed camera designs might happen naturally in many scenes.

2.2 Accidental Pinhole Cameras

Accidental pinhole cameras happen everywhere by the accidental arrangement of surfaces in the world. The images formed are generally too faint and blurry to be noticed, or they are misinterpreted as shadows or inter-reflections. Let's start by showing some examples of accidental pinhole cameras.

One of the most common situations that we often encounter is the pinhole cameras formed by the spacing between the leaves of a tree (e.g. Minnaert 1954). This is

Fig. 4 (**a**) Shows a picture of the floor taken under the shadow of a tree. The pinholes created by the leaves project different copies of the sun on the floor. (**b**) shows a tree inside a corridor near a window produces copies of the scene outside the window. However in this case they are too faint and blurry to be clearly noticed by a person walking by Fig. 10 shows the result of processing this image to increase the contrast

Fig. 5 In this picture, a small cabin in the wall contains a hole pointing downwards. (**a**) The hole acts as a pinhole projecting a green patch on the ceiling. (**b**) View outside the hole. This hole was used as a toilet by the guard of this jail of the XVI century in Pedraza, Spain

illustrated in Fig. 4a showing a picture of the floor taken the shadow of a tree. The tiny holes between the leaves of a tree create a multitude of pinholes. The pinholes created by the leaves project different copies of the sun on the floor. This is something we see often but rarely think about the origin of the bright spots that appear on the ground. In fact, the leaves of a tree create pinholes that produce images in many other situations. In Fig. 4b, a tree inside a corridor near a window produces copies of the scene outside the window. However, in this case, the produced images are too faint and blurry to be clearly noticed by a person walking by.

Figure 5 shows another common situation. Sometimes, small apertures in a scene can project colored lights into walls and ceilings. In this picture, a window contains a hole
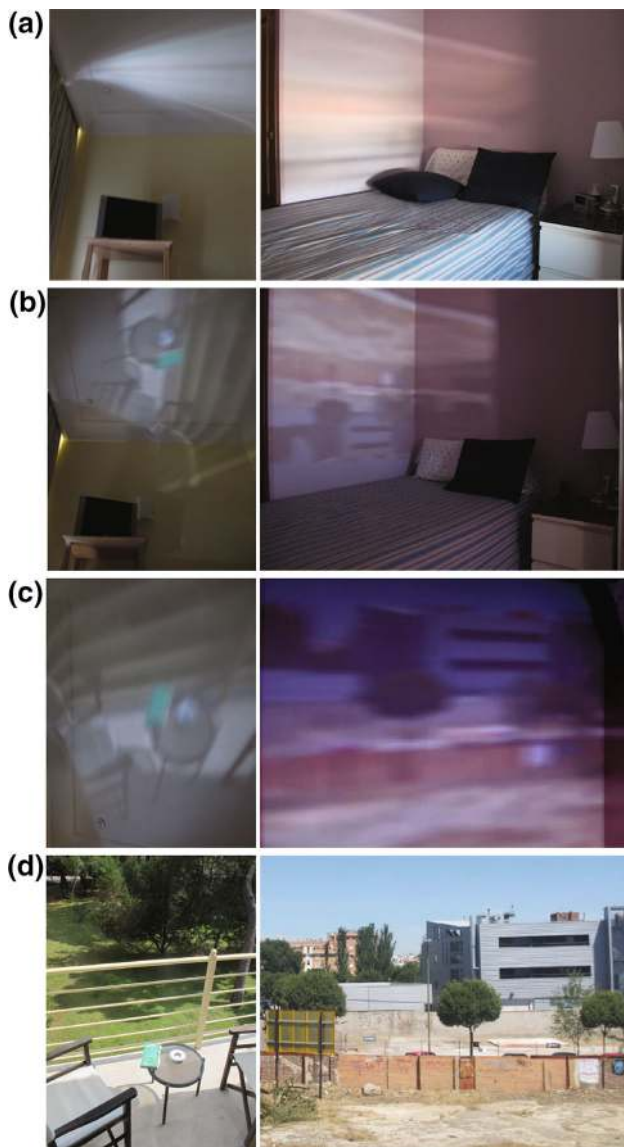
Fig. 6 The *top row* (**a**) shows two different rooms illuminated by exterior light, creating shading patterns within the room. Some of these patterns may look like shadows. The images in (**b**), from the same viewpoints as (**a**), show the effect of closing the windows, leaving only a small aperture, turning the room in a camera obscura. (**c**) Shows those images upside-down, to better reveal the formed image. (**d**) Shows the view from the window to the outside. The shadows on (**a**) are in fact blurred images, not shadows. The room created an accidental camera obscura

Fig. 7 Examples of convolutions by the aperture function. (**a**) Lighting within room shown together with the window opening. (**b**) Lighting from a night scene, and (**c**) the view out the window at night, showing the multiple point sources. The point sources reveal in (**b**) the rectangular convolution kernel of the window aperture. (**d**) Daytime view within the room, and (**e**) the view out the window, which, convolved with the window aperture, yields the projected patterns in (**d**)

ceived at first as shadows or inter-reflections. Figure 6c shows the images re-oriented to allow a better interpretation of the projected image and Fig. 6d shows pictures of what is outside of the window in each case.

Accidental pinhole cameras deviate from ideal pinhole cameras in several ways:

– Large non-circular aperture
– Image is projected on a complex surface far from the ideal white flat lambertian surface.
– Multiple apertures
– Inter reflections (e.g. inside a room the walls will not be black)

To illustrate the image formation process with a room-size example, consider the room shown in Fig. 7a. In this scene, the light illuminating the room enters via a partially open window. In this particular setup, the room will act as a camera obscura with the window acting as the aperture. For simplicity, let's focus on analyzing the image formed on the flat wall opposite to the window (the leftmost wall in Fig. 7a). If the window was a small pinhole, the image projected in the wall

pointing downwards. The hole looks over the ground bellow which is covered by grass and receives direct sunlight. The hole acts as a pinhole projecting a green patch on the ceiling.

Perhaps the most common scenario that creates accidental pinhole cameras is a room with an open window as discussed in Fig. 2. Figure 6a shows two indoor scenes with complex patterns of lights appearing on the walls and ceiling. By transforming each room into a camera obscura, the images appear in focus (Fig. 6b), revealing the origin of what could be per-
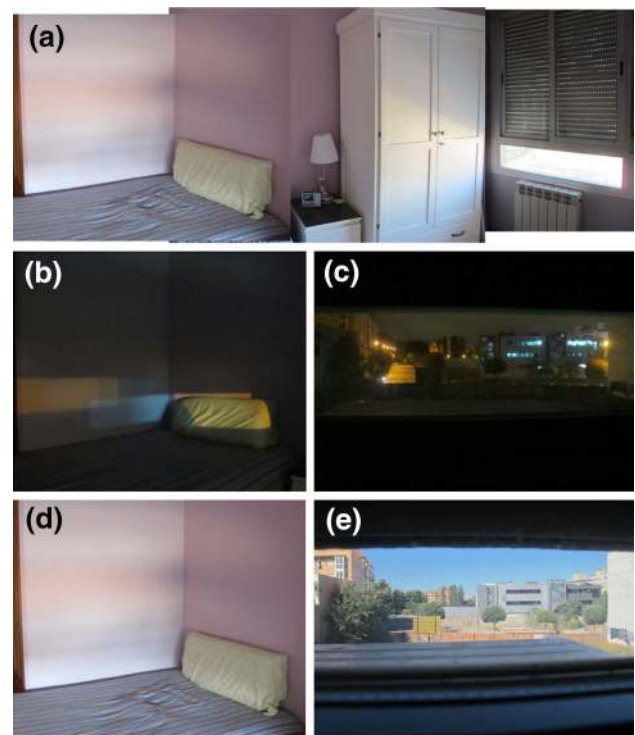
would be a sharp image (as shown in Fig. 6b). Let's denote as $S(x, y)$ the image that would be formed on the wall if the window was an ideal pinhole. As the room deviates from the ideal pinhole camera, the image formed will be different from $S(x, y)$ in several ways. The point spread function produced by the window on the wall, $T(x, y)$, will resemble an horizontally oriented rectangular function. A pinhole camera is obtained when the aperture $T(x, y)$ is sufficiently small to generate a sharp image $I(x, y)$. For a more complete analysis of variations around the pinhole camera we refer to Zomet and Nayar (2006). The resulting image projected on the wall will be the convolution:

$$L(x, y) = T(x, y) * S(x, y) \qquad (1)$$

As the wall will be different from a white lambertian surface, we need to include also albedo variations of the surface where the image is being projected:

$$I(x, y) = \rho(x, y)L(x, y) \qquad (2)$$

Figure 7b, d show two views of the same room under different outdoor illuminations (night time and daylight). At night, illumination sources produce an $S(x, y)$ image that could be approximated by a few delta functions representing the point light sources in the outside scene. Therefore, the image that appears on the wall looks like a few superimposed copies of the window shape (and the coloring indicates which light source is responsible of each copy). Under daylight (Fig. 7d, e), most of the illumination is diffuse, and the resulting image is the convolution of the outdoor scene with the window shape, giving a very blurry image of what is outside. We will show later how this simple model can be used to infer the shape of the window when the window is not visible in the picture.

What we have discussed here is a very simple model that will not account for all the complexities of image formation process and the image hidden inside a room. We have ignored the 3D layout of the scene, variations of the BRDF, inter-reflections (which will be very important as a room is composed of surfaces with different reflectances and colors). Despite its simplicity, this model is useful to suggest successful ways of extracting images of the outside scene.

2.3 Getting a Picture

The images formed by accidental pinhole cameras are blurry and faint, and are generally masked by the overall diffuse illumination and the reflectance of the scene they are projected onto. To increase the contrast of these accidental images we need first to remove from the picture other sources of intensity variation. This problem is generally formulated as finding the intrinsic images (Barrow and Tenenbaum 1978), decomposing the image $I(x, y)$ into a reflectance image $\rho(x, y)$ and an illumination image $L(x, y)$. In the examples in this section we will show that a simple version of the Retinex algorithm (Land and McCann 1971) is quite successful in extracting accidental images from pictures.

There are three main sources of intensity variations superimposed in an accidental camera image:

1. the reflectance image of the interior scene
2. the shading components of the interior scene
3. the projected image of the outside world, blurred by the accidental camera aperture.

Retinex has been used to separate (1) from (2) as in Barrow and Tenenbaum (1978) and in Tappen et al. (2005), but we're using it to separate (3) from the combination of (1) and (2). Retinex works much better for the task of extracting accidental images than to separate (1) from (2), because the accidental camera aperture blurs things so much.

In our setting we are interested in the illumination image $L(x, y)$, removing the effects of the albedo $\rho(x, y)$ of the surface in which the outside image gets projected. Using logarithms, denoted by primes, Eq. 2 becomes:

$$I'(x, y) = \rho'(x, y) + L'(x, y) \qquad (3)$$

Given $I'(x, y)$, our goal is to recover $L'(x, y)$. Land and McCann (Land and McCann 1971) introduced the Retinex algorithm to solve this problem. Since then, there has been a large number of approaches dealing with this problem (e.g. Tappen et al. 2005; Grosse et al. 2009; Barron and Malik 2012). Here we will make use of the same assumption as it was originally proposed by Land and McCann: that the illumination image, $L'(x, y)$, introduces edges in the image that are of lower contrast (and blurrier) than the edges due to the scene reflectance, $\rho'(x, y)$. Although this assumption might work well under direct illumination where strong and sharp shadows appear in the image, it holds true for the situations in which accidental cameras are formed, as the illumination is generally indirect and produces faint variations in the scene.

Retinex works by thresholding the gradients and assigning the gradients below the threshold to the gradients of the illumination image. Here we will use the Canny edge detector (Canny 1986) as a robust thresholding operator as it takes into account not just the local strength of the derivatives but also the continuation of edges in the image. Pixels marked as edges by the Canny edge are more likely to be due to reflectance changes than to variations in the illumination image. We will estimate the gradients of the logarithm of the illumination image as:

$$L'_x(x, y) = I'_x(x, y) \times (1 - E_d(x, y)) \qquad (4)$$
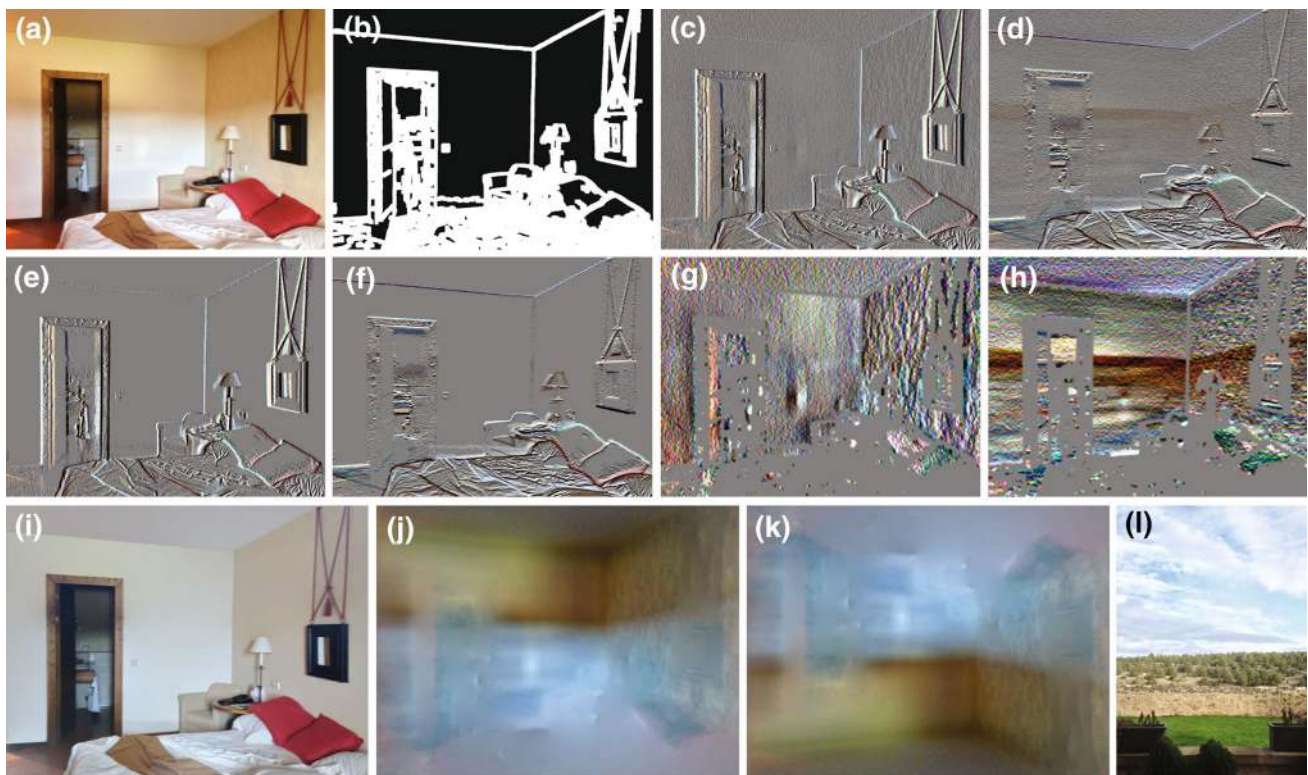$$L'_y(x, y) = I'_y(x, y) \times (1 - E_d(x, y)) \qquad (5)$$

Fig. 8 (**a**) Input image, (**b**) Canny edges, (**c**) $I'_x(x, y)$, (**d**) $I'_y(x, y)$, (**e**) $\rho'_x(x, y)$, (**f**) $\rho'_y(x, y)$, (**g**) $L'_x(x, y)$, (**h**) $L'_y(x, y)$, (**i**) recovered reflectance image, (**j**) recovered illumination image, (**k**) illumination image upside-down, (**l**) view outside of the image

$E_d(x, y)$ is the binary output of the Canny edge detector. The binary mask is made thick by marking pixels that are at a distance of d pixels from an edge. As the illumination image is very faint, it is important to suppress the derivatives due to the albedo that are at some small distance from the detected edges. Once the illumination derivatives are estimated we recover the illumination image that matches those gradients as closely as possible. We use the pseudo-inverse method proposed in Weiss (2001) to integrate the gradient field and to recover the illumination. The method builds the pseudo-inverse of the linear system of equations that computes the derivatives from the illumination image. The pseudo-inverse allows computing the illumination image that minimize the squared error between the observed derivatives and the reconstructed derivatives. Once the illumination image has been estimated, the reflectance image is obtained from Eq. 3.

Figure 8 shows the result of applying Retinex to an input image. Figure 8a shows a picture of a bedroom. The estimated reflectance and illumination images are shown in (i) and (j) respectively. Note that the recovered illumination image has a strong chromatic component. The illumination image is produced by light entering by a window on the opposite wall (not visible in the input image). Therefore, it is an upside-down image of the scene outside the window. Figure 8k shows the upside-down illumination image and Fig. 8l shows the true view outside the window. The illumination image is distorted

due to the room shape but it clearly shows the blue of the sky, and the green patch of the grass on the ground. Figure 9 shows additional results.

As discussed at the beginning of this section, Fig. 4 described how the tiny holes between the leaves of a tree can create a multitude of pinholes. Figure 10 shows the detail from the tree picture shown in Fig. 9. On the wall we can now appreciate that there are multiple repetitions of the blue and orange patches that correspond to the scene outside the window (Fig. 9d).

Unfortunately, the blur factor is generally too large for the images recovered from accidental pinhole cameras to be recognizable. In the next section we introduce another type of accidental camera that can recover, in certain cases, sharper images than the ones obtained with accidental pinhole cameras.

3 Accidental Pinspeck Cameras

Pinhole cameras can be great cameras, but when formed accidentally, the images they create have very poor quality. Here we will discuss pinspeck cameras. Pinspeck cameras are harder to use and less practical than a pinhole camera. However, accidental pinspeck cameras are better and more common than accidental pinhole cameras.
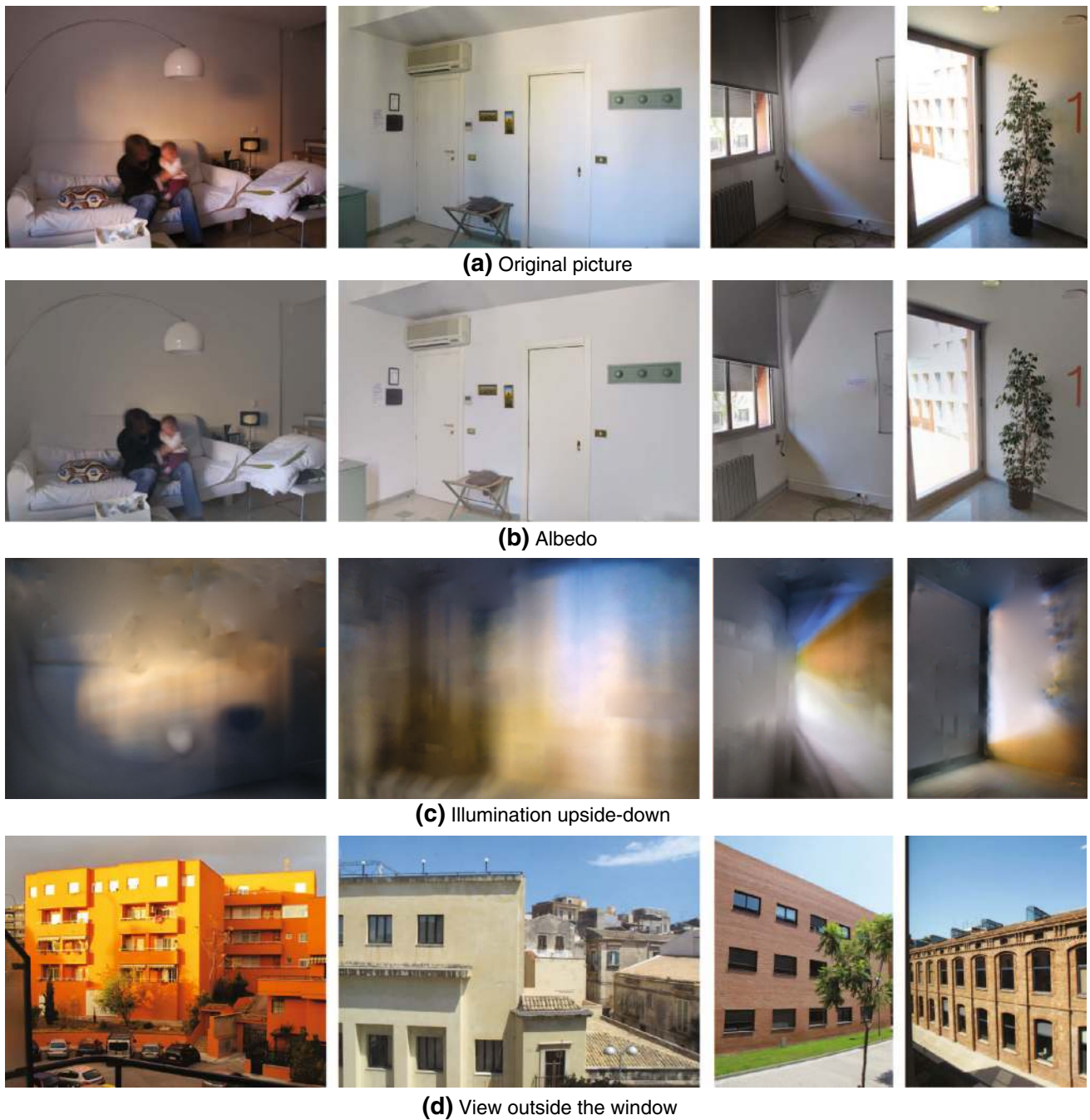
(a) Original picture

(b) Albedo

(c) Illumination upside-down

(d) View outside the window

Fig. 9 Additional results applying Retinex to several images. (**a**) Input images, (**b**) recovered reflectance images, (**c**) recovered illumination images upside-down, and (**d**) view outside of the windows. Note the resemblance between the images in row c and row d (accounting for blurring and projection). The rightmost column shows a special situation in which the recovered image in row c doesn't look like the image in row d. See Fig. 10 and the associated text for an explanation

3.1 Shadows

Under direct sunlight the shadow produced by an object appears as a sharp distorted copy of the object producing it (Fig. 11a) and there seems to be nothing more special about it. The shadow that accompanies us while we walk disappears as soon as we enter under the shadow of a building (Fig. 11b). However, even when there is no apparent shadow around us, we are still blocking some of the light that fills the space producing a very faint shadow on the ground all around us. In fact, by inspecting Fig. 11b it is hard to see any kind of change in the colors and intensities in the ground near the

Fig. 10 The tiny holes between the leaves of a tree can create a multitude of pinholes. After applying the Retinex algorithm, we can now appreciate that there are multiple repetitions of *blue* and *orange* patches corresponding to the scene outside the window (Fig. 9d) on the wall

Fig. 11 A person walking in the street, (**a**) under direct sunlight the person projects a sharp dark shadow. However, (**b**) when there is no direct sunlight, the shadow seems to disappear, but there are still shadows from the indirect illumination. (**c**) Increasing the contrast reveals a colorful shadow

person. But if we crop the region near the feet and increase the contrast we can see that there is a colorful shadow (see Fig. 11c). The shadow is yellow just along the feet and it takes a blue tone right behind the feet.

We will show in the rest of this section that there is indeed a faint shadow and it is strong enough to be detectable. Why is this important? Because a shadow is also a form of accidental image. The shadow of an object is all the light that is missing because of the object's presence in the scene. If we were able to extract the light that is missing (i.e. the difference between when the object is absent from the scene and when the object is present) we would get an image. This difference image would be the negative of the shadow and it will be approximatively equivalent to the image produced by a pinhole camera with a pinhole with the shape of the occluder.

A shadow is not just a dark region around an object. A shadow is the negative picture of the environment around the object producing it. A shadow (or the colored shadows as called by Minnaert (1954) can be seen as the accidental image created by an accidental anti-pinhole camera (or pinspeck camera, Cohen 1982).

3.2 Pinspeck Camera

Pinhole cameras form images by restricting the light rays that arrive to a surface so that each point on a surface gets light from a different direction. However, another way in which rays of light that hit a surface are restricted is when there is an occluder present in the scene. An occluder blocks cer-
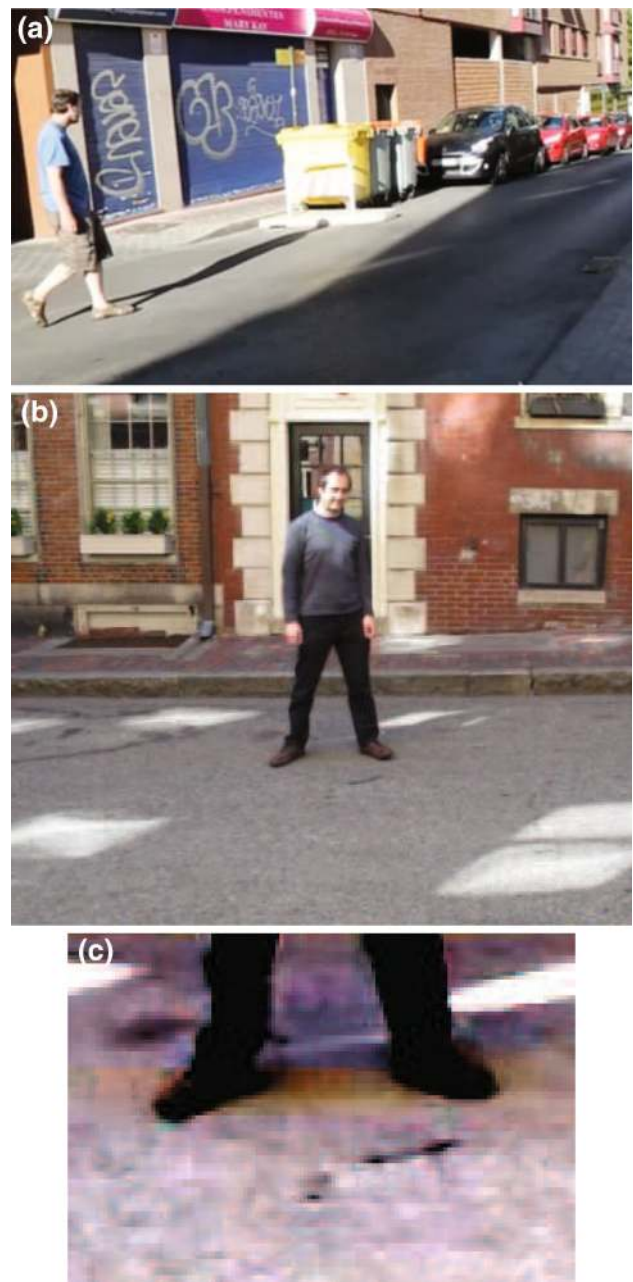
tain of the light rays, producing a diffuse shadow. In the cast shadow, there is more than just the silhouette of the occluder, there is also the negative image of the scene around the occluder. The occluder produces an *anti-pinhole* or *pinspeck* camera.

Pinspeck cameras were proposed by Cohen (1982), and also used before by Zermeno et al. (1978) and Young (1974). Figure 12 illustrates how the pinspeck camera works, as
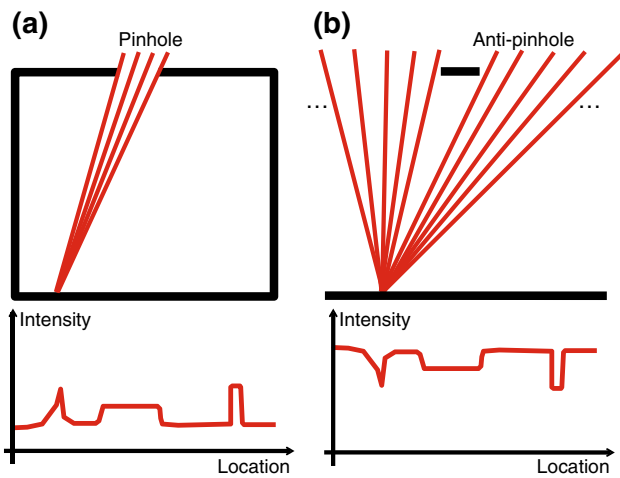
Fig. 12 Illustration of the image formation process for a Pinhole camera (**a**), and a pinspeck camera (**b**). Modified from Cohen (1982)

described by Cohen (1982). In the pinhole camera, a surface inside a box receives light coming from a small aperture. In the pinspeck camera, the box with the hole is replaced by a single occluder. If the occluder size matches the size of the pinhole, the image that gets projected on the surface will have an intensity profile with a bias and reversed with respect to the intensity profile produced by the pinhole camera:

$$L_{occluder}(x, y) = L - L_{pinhole}(x, y), \qquad (6)$$

where L is the overall intensity that would reach each point on the surface if there were no occluder. If the illumination comes from a source infinitely far away, then all the points on the surface will receive the same intensity, L.

As noted by Cohen (1982), there are a number of important differences between the pinspeck and the pinhole camera.

– Bias term L: this term can be quite large in comparison with the light that gets blocked $L_{pinhole}$. Increasing the exposure time will burn the picture. Therefore, in order to improve the signal to noise ratio we need to integrate over multiple pictures.

– Occluder: if the occluder is spherical, the vigneting is reduced as the effective aperture does not change shape when seen from different points on the surface. Therefore, Eq. 6 is just an approximation for the points directly under the occluder.

In the next section we will show that accidental pinspeck cameras are very common.

3.3 Accidental Pinspeck Cameras

Let's first look at a few relaxed pinspeck camera designs. Figure 13 shows some frames of a video showing a ball bouncing. There is no direct sunlight in this corner of the building. Therefore, no shadow is visible. But after close inspection we can see a faint change in the brightness of the walls as the ball gets closer to the wall and ground. In fact, the shadow produced by the ball extends over most of the wall. Note that now L is not constant any more and the surface where the image should be projected is not a white surface. But we can still compute the difference between a frame where the ball is absent and the frames of the video where the ball is present. The resulting difference image corresponds to a picture that one could take if the scene was illuminated only by the light that was blocked by the ball. This is the light produced by a pinhole camera with the pinhole in the location of the ball.

Figure 14 shows a frame upside-down from the processed video from Fig. 13 and compares it with the scene that was in front of the wall. Despite that this relaxed pinspeck camera differs in many ways from the ideal pinspeck camera, it is able to produce a reasonable, albeit blurry, image of the scene surrounding this building corner.

Accidental anti-pinholes differ from ideal anti-pinholes in several aspects:

– Non-spherical (large) occluder.
– The surface has a varying albedo $\rho(x, y)$.
– The bias term L is not constant. This situation is quite common, especially in indoors as we will discuss later.

Fig. 13 Relaxing the anti-pinhole camera. This figure shows some frames of a video showing a ball bouncing and the difference between a frame without ball present and the frames of the video. The difference corresponds to the light that would had been produced by a pinhole camera with the pinhole in the location of the ball. For clarity, the ball is shown as it looks in the original frame
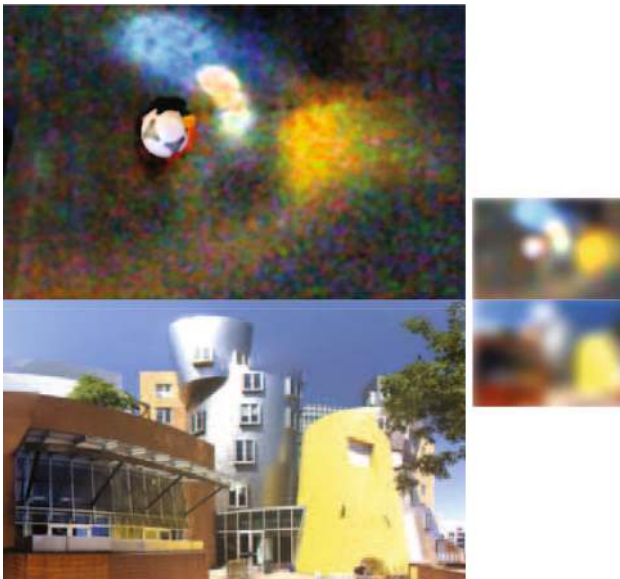
Fig. 14 A frame upside-down from the processed video from Fig. 13 compared with the scene in front of the wall. The right column shows low resolution version of the images in the left column to highlight the similarities between the recovered image (on *top*) and the real scene (*bottom*)

- The scene might have a complicated geometry. For the derivations here we will assume that the portion of the scene of interest is planar.

The goal of the rest of the section is to provide some intuition of how accidental images are formed from accidental pinspeck cameras. We will show how these accidental images can be extracted from sets of pictures or videos. We start by providing an analysis of the image formation process.

If we have an arbitrary scene before the occluder used to form the pinspeck camera is present, we would capture an image that we will call the background image:

$$I_{background}(x, y) = \rho(x, y)L(x, y) \qquad (7)$$

If we had an ideal camera, we would like this image to be constant (with no albedo or illuminations variations). However, the image $I_{background}(x, y)$ will just be a normal picture where variations in intensities are due to both albedo and illumination changes.

If we placed a pinhole to replace the source of illumination, then the image captured would be:

$$I_{pinhole}(x, y) = \rho(x, y)L_{pinhole}(x, y) \qquad (8)$$

and if an occluder appears on the scene, the picture will be:

$$I_{occluder}(x, y) = \rho(x, y)L_{occluder}(x, y) \qquad (9)$$

In this equation we assume that the occluder is not visible in the picture. Note that these three images only differ in the illumination and have the same albedos.

If the pinhole and the occluder have the same silhouette as seen from the surface where the illumination gets projected, then the image captured when there is an occluder can be approximated by:

$$I_{occluder}(x, y) = I_{background}(x, y) - I_{pinhole}(x, y) \qquad (10)$$

and therefore, given two pictures, one of the normal scene and another with the occluder present, we can compute the picture that would had been taken by a pinhole camera with a pinhole equal to the shape of the occluder as:

$$\begin{aligned} I_{pinhole}(x, y) &= I_{background}(x, y) - I_{occluder}(x, y) \\ &= \rho(x, y)\left(L(x, y) - L_{occluder}(x, y)\right) \\ &= \rho(x, y)\left(T_{hole}(x, y) * S(x, y)\right), \qquad (11) \end{aligned}$$

where $T_{hole}(x, y)$ is related to the occluder silhouette and $\rho(x, y)$ is the surface albedo.

If $L(x, y)$ is constant, then we can remove the unknown albedo by using the ratio of the image with the occluder and the image without it:

$$L_{pinhole}(x, y)/L = 1 - \frac{I_{occluder}(x, y)}{I_{background}(x, y)} \qquad (12)$$

However, $L(x, y)$ is rarely constant in indoor scenes and computing ratios will not extract the desired image.

Figire 15 shows a few frames of a video captured at the same scene as in Fig. 13 but with a person walking instead of the bouncing ball. In order to apply Eq. 11 we first compute a background image by averaging the first 50 frames of the video before the person entered the view. Then, we compute the difference between that background image and all the frames of the video to obtain a new video showing only the scene as if it was illuminated by the light that was blocked by the person. Three frames of the resulting video are shown in Fig. 15.

We will study next typical situations in which accidental pinspeck cameras occur.

3.4 Shadows in Rooms

The indoors provide many opportunities for creating accidental cameras. As discussed in Sect. 2, a room with an open window can become an accidental pinhole camera. In Sect. 2 we showed how we could use Retinex in order to estimate the illumination image $L_{pinhole}(x, y)$. Despite that we can recover images revealing some features of the scene outside the room (Fig. 9), the images generally reveal only a few color patches and are too blurry to be recognizable.

Let's now imagine that we have access to several images of the room, or a video, where a person is moving inside
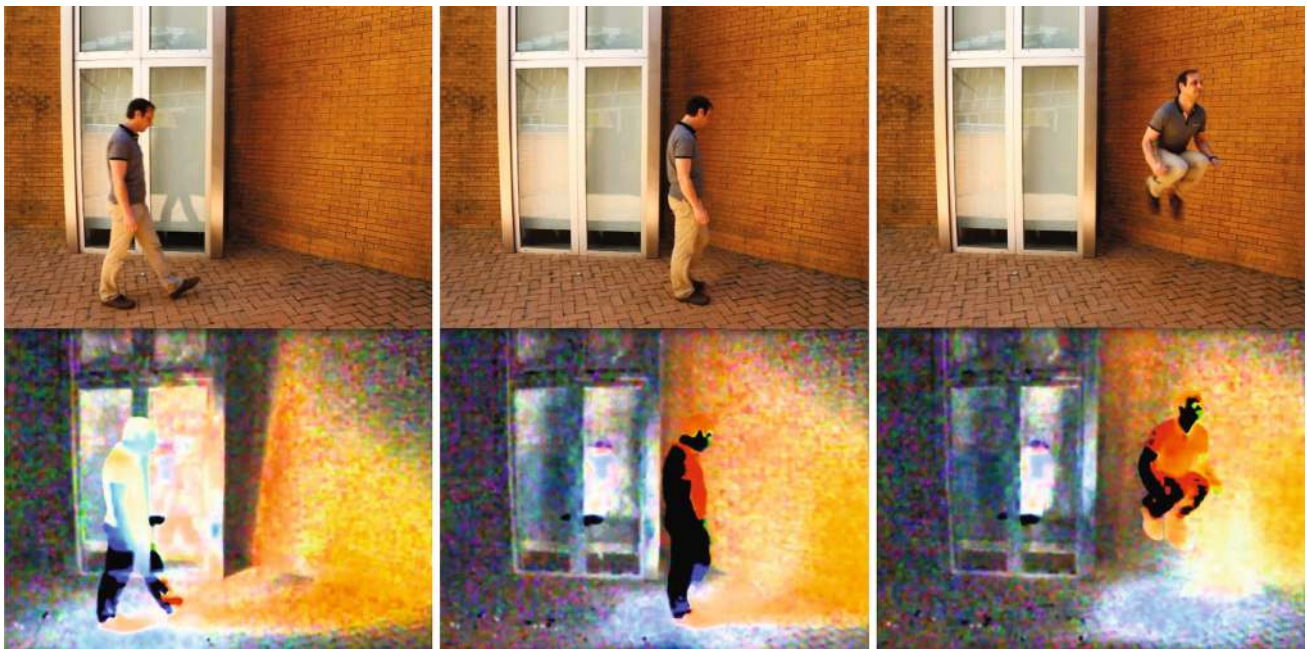
Fig. 15 Relaxing the anti-pinhole camera. Compare with Fig. 13. The man forms a fairly large occluder, leading to a blurry pin speck camera image, in contrast with that of the ball, in Fig. 13. At the far right, the man tries to become a better pinhole, which helps a little

Fig. 16 Three frames from a video of a person walking inside a room. *Top row* shows the three unprocessed frames, and the *bottom row* shows the difference of multi-frame average centered on current frame from a multi-frame average of the background. (**a**) One of the first frames in the video. (**b**) A person inside the room blocks some of the light entering the window and produces a colorful shadow (**c**) and the person is not visible anymore, but now a faint but sharp image gets projected onto the wall. In this last frame, the person is very close to the window producing a better accidental camera

the room. As the person moves, it will be blocking some of the ambient light. The person will behave as an accidental pinspeck camera. To extract a picture from this accidental pinspeck camera inside the room we will apply Eq. 11. First, we use 50 frames from the sequence to compute $I_{background}(x, y)$. Then, we subtract all the frames of the video from that background image. Figure 16 shows three frames from the video. The first frame (Fig. 16a) corre-sponds to the beginning of the video and it is very similar to the background image as the person has not entered the scene yet. Therefore, applying Eq. 11 to this frame results mostly in noise. Later in the video, a person enters in the room (Fig. 16b) blocking some of the light entering the window and producing a colorful shadow. However, the obtained dif-ference image from Eq. 11 is not much better than the image obtained with the Retinex algorithm. However, later on the
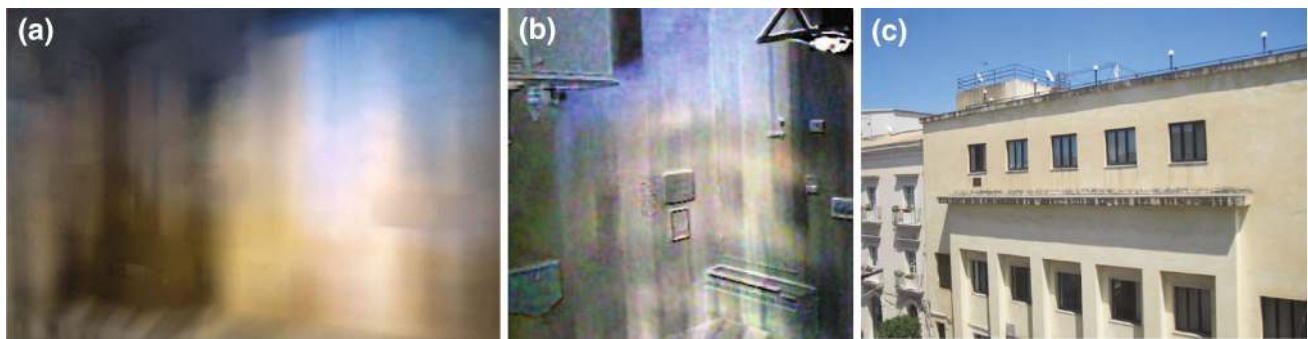
Fig. 17 Comparison between the accidental pinhole and the accidental pinspeck cameras. (**a**) Output of Retinex on a single frame from Sect. 2.3, designed to extract pinhole camera image. (**b**) Output of the accidental pinspeck camera (selected frame), and (**c**) true view outside the window. (**a**), (**b**) Upside-down so that they can be compared easily with (**c**). As is often the case, this pinspeck camera image is noisier, but sharper, than the related pinhole camera image
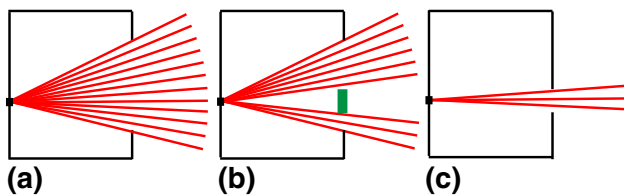

Fig. 18 (**a**) Room with a big aperture (too large to produce a sharp image), (**b**) aperture with an occluder, (**c**) difference between the two light fields, revealing just the light rays striking the small occluder

Fig. 19 (**a**) Window, (**b**) window with an occluder, (**c**) view of the wall opposite to the window when no occluder is present, (**d**) view of the wall with the occluder present

video a faint but sharp image gets projected onto the wall when applying Eq. 11. In that frame the person is not visible within the picture, but it is still blocking part of the light producing now a much better accidental camera than the one formed by the room alone. Figure 17 compares the images obtained with the accidental pinhole camera (Fig. 17a) and the picture obtained from the video (Fig. 17b). Figure 17c shows the view outside the window. The building is now recognizable in Fig. 17b. What has happened here?

As the person was walking inside the room eventually he passed in front of the window. At that moment, the occluder became the size of the intersection between the person and the window, which is much smaller than the person or the window. This scenario is illustrated in Fig. 18. Figure 18 shows how an occluder produces light rays complementary to that of a small aperture with the size of the occluder. Figure 18a shows the rays inside a room that enter via a window. The figure shows all the light rays that hit a point inside the room (in this drawing we assume that there are no interreflections and that all the light comes from the outside). Figure 18b shows the light rays when there is an occluder placed near the window. The difference between the two light fields is illustrated in Fig. 18c. The intersection between the person and the window creates a new equivalent occluder:

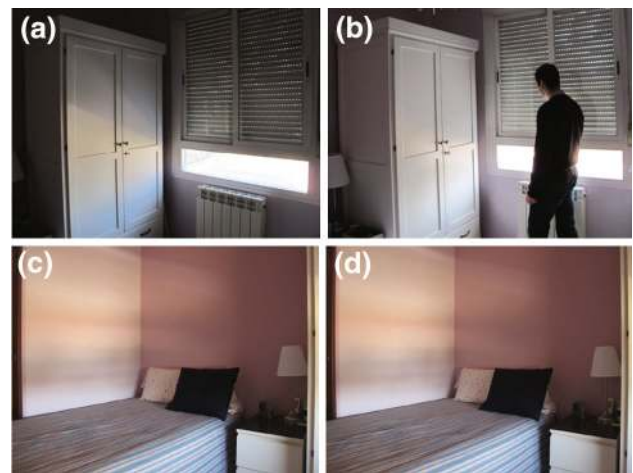$$T_{hole}(x, y) = T_{person}(x, y) \times T_{window}(x, y) \qquad (13)$$

and, therefore:

$$\begin{aligned} I_{window}(x, y) &- I_{occluded-window}(x, y) \\ &= \rho(x, y) \left(T_{hole}(x, y) * S(x, y) \right) \end{aligned} \qquad (14)$$

As $T_{hole}(x, y)$ can be now small, the produced image becomes sharper than with the image produced just by the window alone.

Figure 19 shows another example showing pictures of the window to illustrate how the person is located with respect to the window (Fig. 19a, b). All the illumination in the room is coming via the window. Figure 19c, d show the corresponding pictures on showing the wall in front of the window. There is a very small difference between images (c) and (d), but that difference carries information about the scene that can be seen through the window. Note in this case that Fig. 19c corresponds to $I_{background}(x, y)$ in Eq. 11. In this case $L(x, y)$ is clearly not constant as the illumination in the scene that
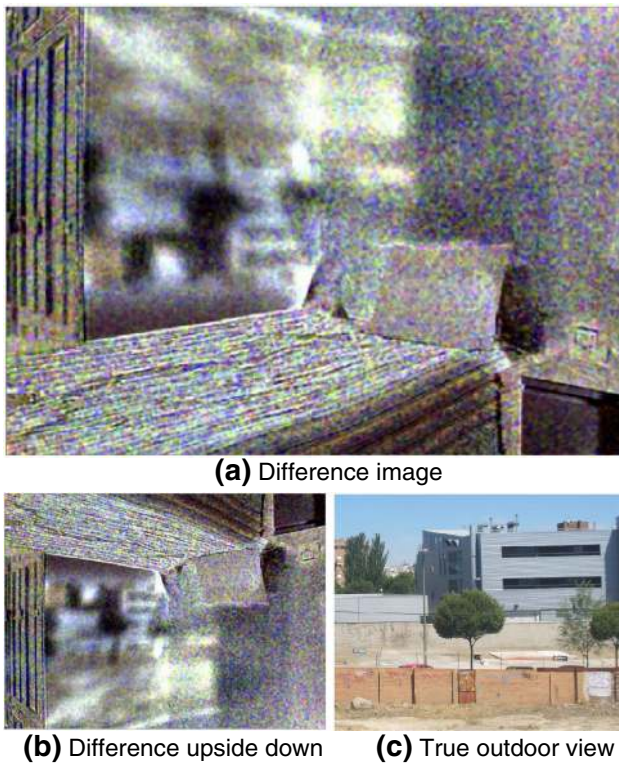
(a) Difference image

(b) Difference upside down   **(c)** True outdoor view

Fig. 20 (**a**) Difference image (Fig. 19c minus Fig. 19d). (**b**) Difference upside-down. (**c**) True outside scene

projects to the wall is already the result of an accidental pinhole camera. Therefore, we can not use ratios to remove the effect of albedo variations in the scene.

In order to recover the image that would have been produced by a pinhole with the shape of the intersection between the person and the window we need to subtract two images—the image with the occluder (Fig. 19d) from the image without it (Fig. 19c).

Figure 20a shows the difference image obtained by subtracting Fig. 19d from Fig. 19c. In the difference image we can see an increased noise level because we are subtracting two very similar images. But we can also appreciate that a pattern, hidden in the images from Fig. 19, is revealed. This pattern is a picture of what is outside the room as it would had been obtained by the light entering the room by an aperture of the size and shape of the occluder. By making the occluder smaller we can get a sharper image, but at a cost of increased noise.

Figure 21 shows the input video and the difference between the background image and the input video. The first frame is only noise, but as the person moves we can see how the wall reveals a picture. As the person moves, the occluder produces a pinhole camera with the pinhole in different locations. This produces a translation on the picture that appears on the wall. These translated copies of the image contain disparity information and could be used to recover the 3D structure if the noise is low enough.

3.5 Limitations

The inverse pinhole has two limitations over traditional pinhole cameras. The first is that it requires at least two images or a video because we need to extract a reference background. The second limitation relates to signal to noise ratio. If the picture had no noise and unlimited precision, it would be possible to extract a perfect sharp image (after deblurring) from the inverse pinhole. In general, to improve the signal to noise ratio (SNR), traditional pinhole cameras require increasing the sensitivity of the light sensor or using long exposures in order to capture enough light. In inverse pinhole cameras the signal to noise ratio decreases when the background illumination increases with respect to the amount of light blocked

Fig. 21 *Top row* Input sequence (a person walks inside a room moving toward and from a window not visible in the movie), *bottom row*) difference between reference image (first frame of the video)

and each frame. The difference creates an approximation to a camera obscura with an aperture that moves as the occluder moves inside the room
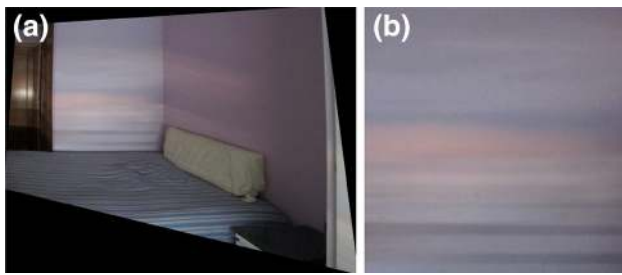
Fig. 22 (**a**) Rectified image, and (**b**) crop and rectified wall from Figs. 7a and 21

by the occluder. If the input is a video, then temporal integration can improve the signal to noise ratio.

While there are many causes of noise in images (Liu et al. 2008), if we assume just Poisson noise, proportional to the square root of the light intensity, we can calculate the SNR of the computed image, limited by the discrete nature of light. Let A be the area of an aperture, $A = \int T(x)dx$. The SNR of the unoccluded photo will be proportional to $\sqrt{A_{window}}$. The signal of the difference image is proportional to $A_{occluder}$, while its noise is proportional to $\sqrt{A_{window}}$, giving an SNR of $\frac{A_{occluder}}{\sqrt{A_{window}}}$. Thus the SNR of the accidental image is reduced from that of the original image by a factor of $\frac{A_{occluder}}{A_{window}}$. Specifics of the sensor noise will reduce the SNR further from that fundamental limit. Therefore, this method will work best when the light entering the room comes from a small window or a partially closed window. In such a case, the ratio between the image without the occluder and the difference image will have similar intensity magnitudes. There are also other sources of noise, like interreflections coming from the walls and other objects.

Despite these limitations, accidental pinspeck cameras might be used to reveal information about the scene surrounding a picture not available by other means. We will discuss some applications in Sect. 4. As discussed before, in order to get a sharp image when using a pinhole camera, we need to make a small aperture. This is unlikely to happen accidentally. However, it is more common to have small occluders entering a scene.

3.6 Calibration

One important source of distortion comes from the relative orientation between the camera and the surface (or surfaces) in which the image is projected. Figure 22 shows how the wall from Figs. 7a and 21 is corrected by finding the homography between the wall and the camera. This can be done by using single view metrology (e.g. Criminisi et al. 2000). This correction is important in order to use the images to infer the window shape, in Sect. 4.3.

We have the additional difficulty of finding the reference image (the image without the occluder). If the input is on

video, one way of deciding which frame can be used as reference is to select the frame with highest intensity (as the occluder will reduce the amount of light entering into the scene). Another possibility is to use multiple frames as reference and select the one providing more visually interpretable results.

4 Applications of Accidental Cameras

In this section we will discuss several applications of accidental cameras.

4.1 Seeing What is Outside the Room

Paraphrasing Abelardo Morell (1995), "a camera obscura has been used ... to bring images from the outside into a darkened room". As shown in Sect. 3.2, in certain conditions, we can use the diffuse shadows produced by occluders near a window to extract a picture of what is outside of the room and we have shown numerous examples of accidental pinhole and pinspeck cameras inside rooms. Figure 23 shows a different example inside a bedroom.

As discussed before, to extract accidental images we need to find the reference image to apply eq. 11. In the case of Fig. 21 we used the average of the first 50 frames of the video. But nothing prevents us from using different reference images. Using different reference images might actually create new opportunities to reveal accidental images. This is illustrated in Fig. 24.

Figure 24 shows a few frames from a video in which a wall and a window are visible. A person walks in the room and stands near the window. In the first frame Fig. 24a, the person is not near the window and it can be used as reference frame. If we subtract from this picture the one from frame Fig. 24b, we obtain the image shown in Fig. 24d which reveals the scene outside the window. The scene is still quite blurred. However, if we continue watching the video, there is a portion of the video where the person is standing near the window and just moves one hand (Fig. 24c). If we use now as reference Fig. 24b and we subtract Fig. 24c, this will correspond to an accidental camera with a pinhole equal to the size of the intersection between the window and the arm. That is a much smaller occluder than the one obtained before. The result Fig. 24g. This is a sharper image (although noisier) than the one obtained before. Figure 24f–h compare the two accidental images with the true view outside the window.

4.2 Seeing Light Sources

In indoor settings, most of the illumination is dominated by direct lighting. Due to the large ratio between direct and indirect illumination when there are direct light sources, shadows

Fig. 23 Finding a picture of what is outside a room (**d**) from two pictures (**a**) and (**b**). The true view (**e**) is shown for comparison with the recovered image (**d**)
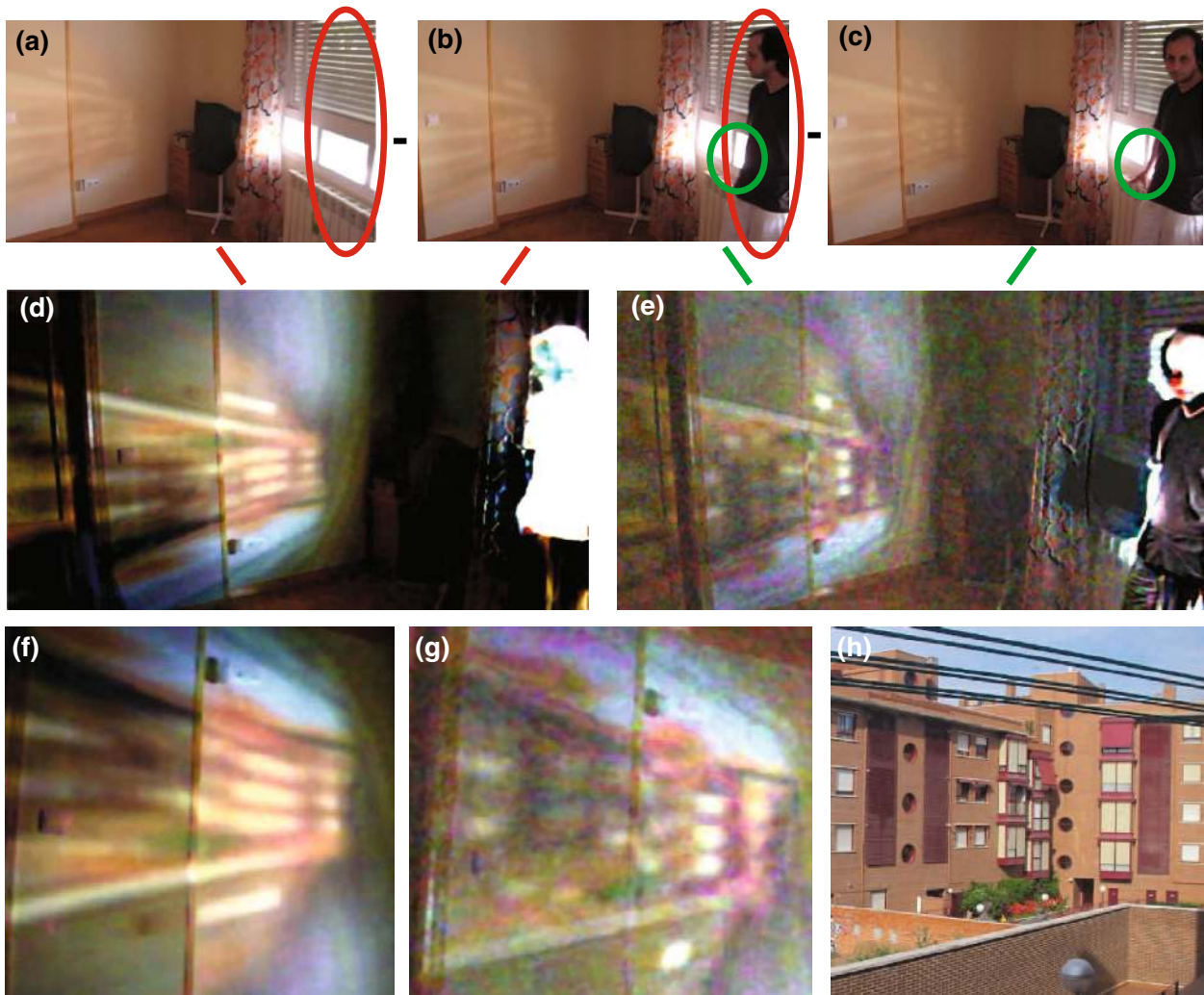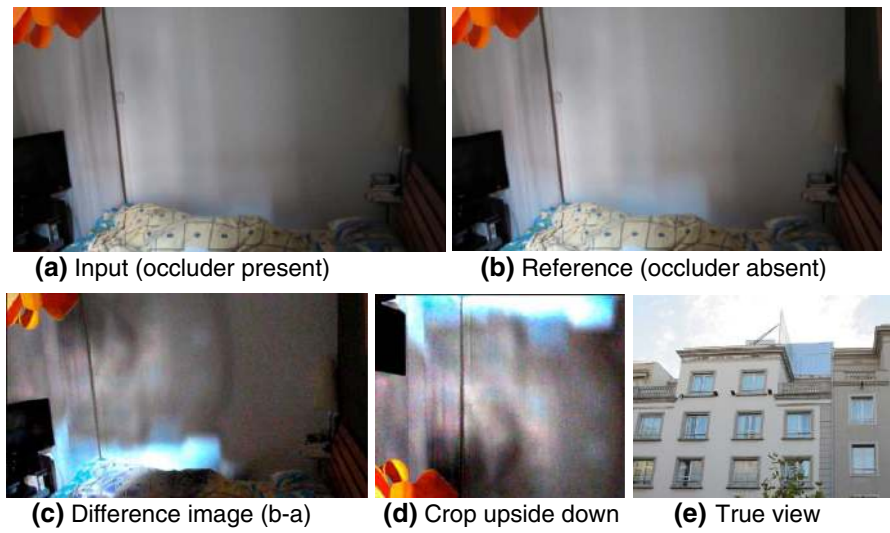

(a) Input (occluder present)   **(b)** Reference (occluder absent)

(c) Difference image (b-a)   **(d)** Crop upside down   **(e)** True view

Fig. 24 Looking for different accidental images within a sequence. (**a**)–(**c**) Show three frames of a long video. (**d**), (**e**) Show two different accidental images using different reference images. (**f**)–(**h**) Comparison of the accidental images with the true view outside the window. Notice that (**g**), taken using a smaller occluder, is sharper, but noisier
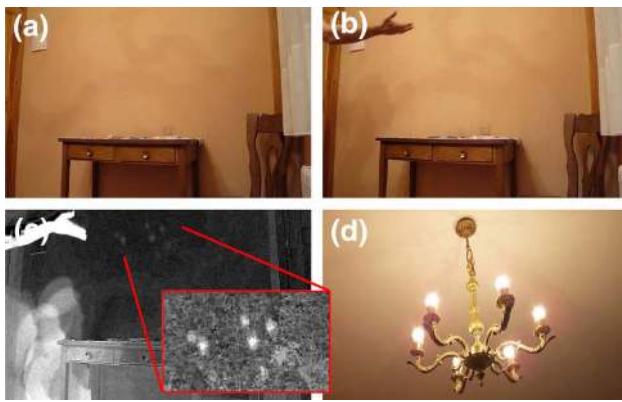
Fig. 25 (**a**) Reference image, and (**b**) image with an occluder producing a faint shadow on the wall. There are two main occluders: a hand and a ball. The ball is already outside of the frame of the picture. (**c**) Difference image. The shadow reveals a person throwing a ball. The ball acts as a pinhole camera and produces a clearer picture of the light sources. (**d**) Picture of the lamp illuminating the scene (ground truth)

can only be used to recover the light sources. If the signal to noise ratio were sufficiently large, it could be possible to get a picture of the rest of the scene. Figure 25 shows an example. In Fig. 25 a ball produces a shadow that can be used to extract a picture of the lamp in the ceiling.

4.3 Seeing the Shape of the Window

Figure 26 shows a series of pictures taken in two different rooms with windows closed by different amounts and with different window shapes. As the window closes, the pattern of illumination inside the room changes. Note that when there is diffuse illumination coming from the outside, the window shape is not clearly visible on the wall. This is clearly illustrated on Fig. 7. Figure 7 shows that when there are point light sources outside, the window shape appears clearly projected onto the wall. However, with more general outdoor scenes, the window shape is not visible directly. However the window shape has a strong influence on the blur and gradient statistics of the pattern projected onto the wall.

As discussed in Sect. 2.1, the pattern of intensities on the wall corresponds to a convolution between the window shape and the sharp image that would be generated if the window was a perfect pinhole. Therefore, the shape of the window modifies the statistics of the intensities seeing on the wall just as a blur kernel changes the statistics of a sharp image. This motivates using algorithms from image deblurring to infer the shape of the window. The shape of the window can be estimated similarly to how the blur kernel produced by motion blur is identified in the image deblurring problem (e.g. Krishnan et al. (2011)).

Figure 26 shows the estimated window shapes using the algorithm from Krishnan et al. (2011). The input to the algorithm are the images from Fig. 26c, g and the output are the

window shapes shown in Fig. 26d, h. The method shows how the kernel gets narrower as the window is closed and it also correctly finds the orientation of the window. It fails only when the window is very open as the pattern of intensities is too blurry, providing very little information.

Finding the light sources, window shape and the scene outside a picture could be used in computer graphics to provide a better model of the light rays in the scene to render synthetic objects that will be inserted inside the picture.

4.4 Seeing the Illumination Map in an Outdoor Scene

Any object in a scene is blocking some light and, effectively behaving like an accidental pinspeck camera taking a picture of its surrounding. In particular, a person walking in the street projects a shadow and acts like an accidental pinspeck camera. In this case the occluder is very large and with a shape very different from a sphere.

As shown in Fig. 11, the shadow around a person can be very colorful. If we have two pictures, one without the person and another with the person, taking the difference between them (Eq. 11) reveals the colors of the scene around the person as shown in Fig. 27a. We can see that the yellow shadow in Fig. 11 corresponded in fact to the blue of the sky right above the person, and the blueish shadow behind it corresponded to a yellow reflection coming from a building in front of the person not visible in the picture. Figure 27b shows the same street but on a cloudy day. Now the colorful shadow has been replaced by a gray shadow. Without strong first-bounce-from-sun lighting, the shadow only shows the gray sky.

Figure 28 shows five frames from a video in which a person is walking in the street. In the first frame from Fig. 28, the person is in a region of the scene where there is direct sunlight. The person creates a sharp image (which is just a picture of the sun projected on the ground and deformed by the person shape and the scene geometry). However, as soon as the person enters the region of the scene that is under the shadow of a building, the shadow becomes faint and increasing the contrast reveals the colors of the scene around the person. In these results the background image is computed as the average of the first 50 frames from the video.

If we know the 3D geometry of the scene and the location of the occluder, then we can infer where the light rays that contribute to the shadow come from and we could reconstruct the scene around the person and outside of the picture frame. This is illustrated in Fig. 29. Figure 29a shows one frame of a sequence with a person walking. Figure 29b shows the background image (computed as the median of all the frames in the video), and Fig. 29c shows the difference (b)–(a), which is the negative of the shadow. In order to recover the 3D geometry we use single view metrology. We use LabelMe 3D which allows recovering metric 3D from

Fig. 26 (**a**, **e**) Window (ground truth), (**b**, **f**) picture of the room, (**c**, **g**) warped and cropped wall region (input to the estimation), and (**d**, **h**) estimated window shape (the estimated shape is quite robust to the size of the estimated kernel size). Note that the kernel estimation algorithm infers the qualitative size and shape of the window apertures in most cases
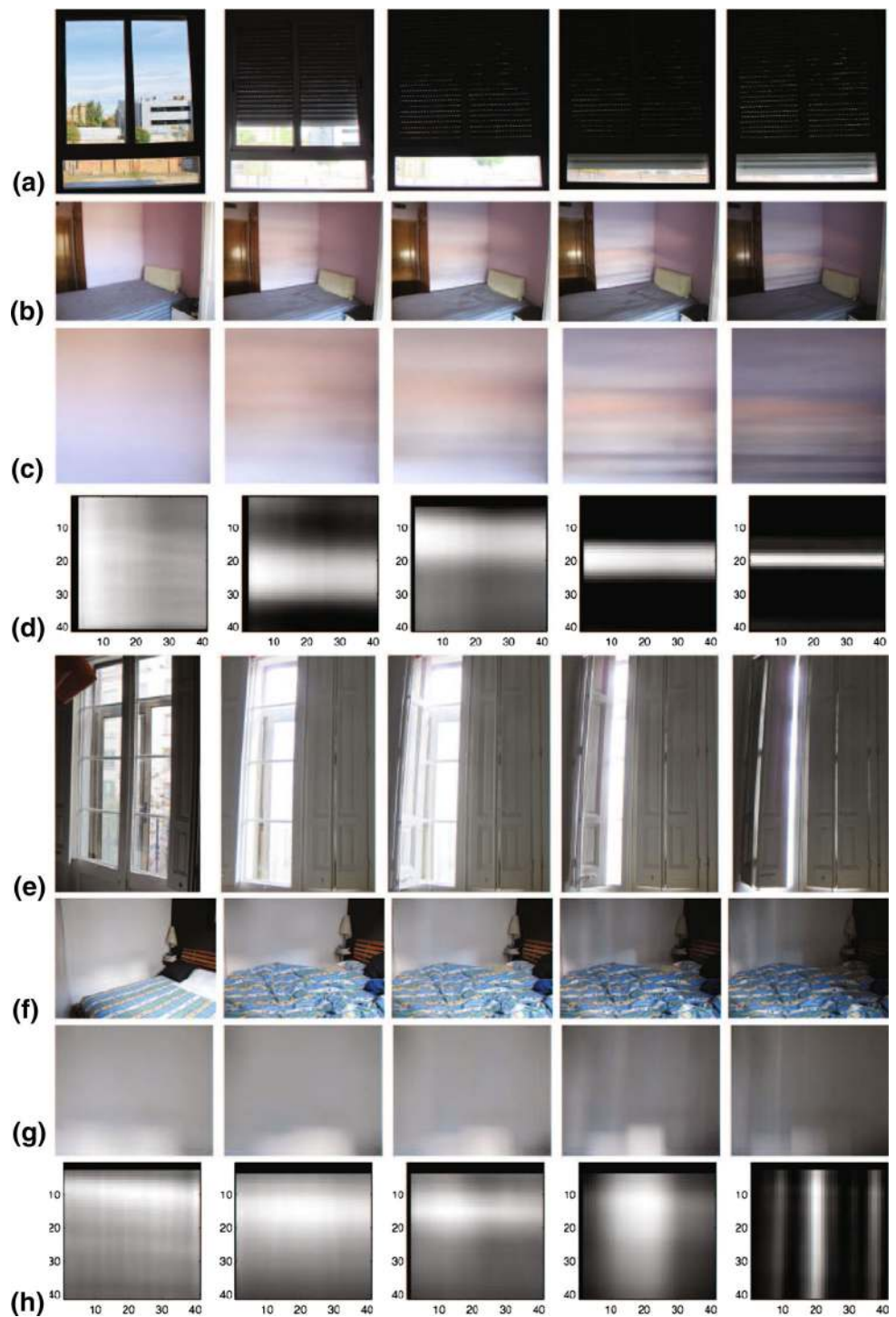
object annotations (Russell and Torralba 2009). The recovered 3D scene is shown in Fig. 29d. Figure 29e shows the panoramic image reconstructed only from the information directly available from the input Fig. 29a. Pixels not directly visible in the input picture as marked black. Figure 29f shows the recovered panorama using the shadow of the person and

Fig. 29g shows a crop of the panorama corresponding to the central region. The yellow region visible in Fig. 29g is in fact a building with a yellow facade. Figure 29h which shows the full scene for comparison. Note that the shadow projected on the wall on the left side of the picture provides information about the right side of the scene not visible inside the picture.

Fig. 27 The colors of shadows on sunny (**a**) and cloudy (**b**) days. The image (**a**) shows the scene from Fig. 11 but now showing the result of applying Eq. 11. (**b**) Shows the same scene on a cloudy day. Now the shadow appears gray

4.5 Accidental Pinholes and Pinspecks Everywhere

Any time an object moves in a video it is creating accidental images. As an object moves, the light rays that reach different parts of the scene change. Most of the times those changes are

very faint and remain unnoticed, or just create sharp shadows. But in some situations, the signal to noise ratio is enough to extract from a video the hidden accidental images formed.

An illustration of how a moving object creates accidental pinhole and pinspeck cameras is shown in Fig. 30. In this video, a person is sitting in front of a computer and moving his hand. Behind the person there is a white wall that receives some of the light coming from the computer screen. As the person moves, there are some changes in the light that reaches the wall. By appropriately choosing which frames need to be subtracted, one can produce the effect of an accidental pinspeck being placed between the screen and the wall. This accidental pinspeck will project a picture of the screen on the wall.

When an object is moving, choosing the best reference frame might be hard. A simple technique that can be applied is to compute temporal derivatives. In order to process the video, we created another video by computing the difference between one frame and the frame two seconds before. The resulting video was temporally blurred by averaging over blocks of ten frames in order to improve the signal to noise ratio. Once the video is processed it has to be inspected to identify which frames produce the best accidental images. Exploring carefully a video can be time consuming and

Fig. 28 Walking on the street. *Shadows* from indirect lighting can be colorful, due the colors of the sky and buildings around the person
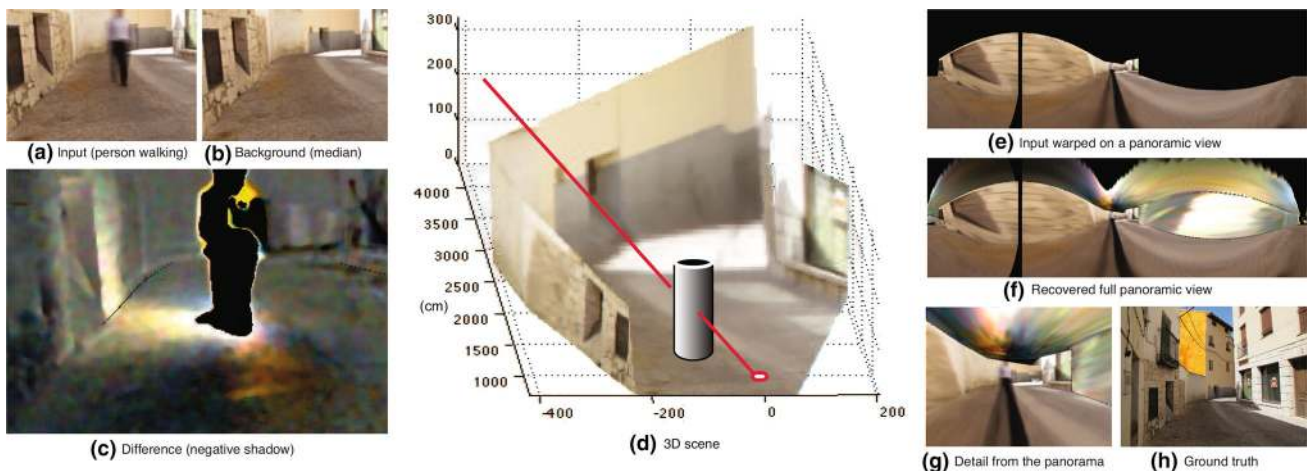

Fig. 29 A person walking in the street projects a complex shadow containing information about the full illumination map outside the picture frame. This figure illustrates how to use the shadow projected by a person (**c**) to recover a panoramic view of the scene outside the picture frame (**g**)

Fig. 30 Accidental pinholes and pinspecks cameras can be generated as an object moves or deforms. (**a**, **b**) Show two frames of a video. (**c**) Difference image revealing a pattern projected on the wall. (**d**) Some of the resulting images formed on the wall compared to the actual image that was shown on the computer screen
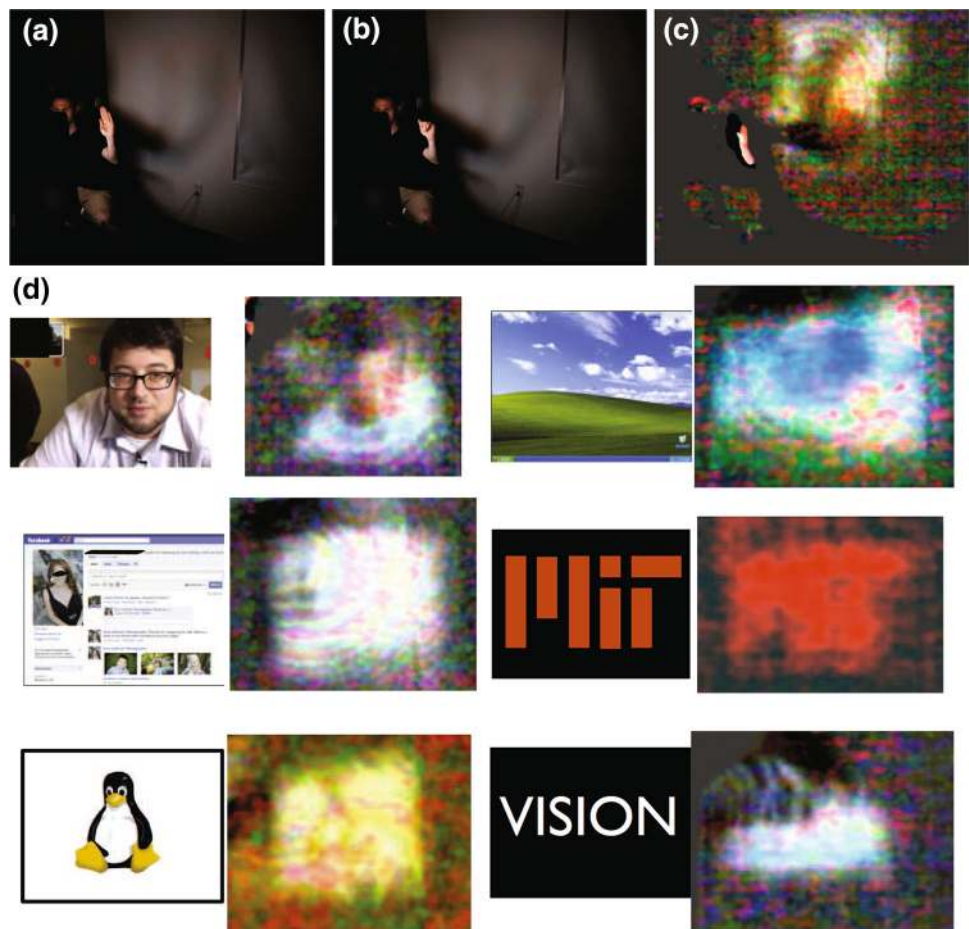
it might require exploring different time intervals to compute derivatives, or chose among different possible reference images.

Figure 30a, b show two selected frames of the video and Fig. 30c shows the difference. We can see that a blurry pattern is projected on the wall behind. That pattern is an upside-down view of the image shown in the screen. Figure 30d shows several examples of what was shown in the screen and a selected frame from the processed video. Despite that the images have low quality they are an example of accidental images formed by objects in the middle of a room.

5 Conclusion

We have described and shown "accidental" images that are sometimes found in scenes. These images can either be direct or processed from several images to exploit "inverse pinholes". These images (a) explain illumination variations that would otherwise be incorrectly attributed to shadows, can reveal (b) the lighting conditions outside the interior scene, or (c) the view outside a room, or (d) the shape of the light aperture into the room, and (e) the illumination map in an outdoor scene. While accidental images are inherently low signal-to-noise images, or are blurry, understanding them is required for a complete understanding of the photometry of many images. Accidental images can reveal parts of the scene that were not inside the photograph or video and can have applications in forensics (O'Brien and Farid 2012).

References

Adelson, E. H., & Wang, J. Y. (1992). Single lens stereo with a plenoptic camera. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, *14*(2), 99–106.

Baker, S., & Nayar, S. (1999). A theory of single-viewpoint catadioptric image formation. *International Journal on Computer Vision*, *35*(2), 175–196.

Barron, J. T., Malik, J. (2012). Shape, albedo, and illumination from a single image of an unknown object. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

Barrow, H. G., Tenenbaum, J. M. (1978). Recovering intrinsic scene characteristics from images. Tech. Rep. 157, AI Center, SRI International, 333 Ravenswood Ave., Menlo Park, CA 94025.

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

Cohen, A. L. (1982). Anti-pinhole imaging. *Optical Acta*, *29*(1), 63–67.

Criminisi, A., Reid, I., & Zisserman, A. (2000). Single view metrology. *International Journal on Computer Vision*, *40*(2), 123–148.

Grosse, R., Johnson, M. K., Adelson, E. H., Freeman, W. T. (2009). Ground-truth dataset and baseline evaluations for intrinsic image algorithms. In *International Conference on Computer Vision*, pp 2335–2342.

Hasinoff, S. W., Levin, A., Goode, P. R., Freeman, W. T. (2011). Diffuse reflectance imaging with astronomical applications. In *IEEE International Conference on Computer Vision*.

Krishnan, D., Tay, T., Fergus, R. (2011). Blind deconvolution using a normalized sparsity measure. In *IEEE Conference on Computer Vision and Pattern Recognition*.

Land, E. H., & McCann, J. J. (1971). Lightness and retinex theory. *Journal of the Optical Society of America*, *61*(1), 1–11.

Levin, A., Fergus, R., Durand, F., Freeman, W. T. (2007). Image and depth from a conventional camera with a coded aperture. In *Proceedings of the SIGGRAPH, ACM Trans On Graphics*, 2007

Liu, C., Szeliski, R., Kang, S. B., Zitnick, C. L., & Freeman, W. T. (2008). Automatic estimation and removal of noise from a single image. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, *30*(2), 299–314.

Minnaert, M. (1954). *The nature of light and color in the open air*. New York: Dover.

Morell, A. (1995). *A camera in a room*. Washington, D.C.: Smithsonian Institution Press.

Nayar, S., Krishnan, G., Grossberg, M., Raskar, R. (2006). Fast separation of direct and global components of a scene using high frequency illumination. In *Proceedings of the SIGGRAPH, ACM Trans On Graphics*, 2006.

Nishino, K., & Nayar, S. K. (2006). Corneal imaging system: Environment from eyes. *International Journal on Computer Vision*, *70*, 23–40.

O'Brien, J., & Farid, H. (2012). Exposing photo manipulation with inconsistent reflections. *ACM Transactions on Graphics*, *31*(1), 1–11.

Russell, B. C., Torralba, A. (2009). Building a database of 3d scenes from user annotations. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp 2711–2718.

Tappen, M. F., Freeman, W. T., & Adelson, E. H. (2005). Recovering intrinsic images from a single image. *IEEE transactions on Pattern Analysis and Machine Intelligence*, *27*(9), 1459–1472.

Torralba, A., Freeman, W. (2012). Accidental pinhole and pinspeck cameras: Revealing the scene outside the picture. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012, pp 374–381.

Weiss, Y. (2001). Deriving intrinsic images from image sequences. In *ICCV*, pp 68–75.

Young, A. T. (1974). Television photometry: The mariner 9 experience. *Icarus*, *21*(3), 262–282.

Zermeno, A., Marsh, L.M., Hevesi, J.M. (1978). Imaging by point absorption of radiation. U.S. patent 4,085,324.

Zomet, A., Nayar, S. K. (2006). Lensless imaging with a controllable aperture. In *IEEE Conference on Computer Vision and Pattern Recognition*.