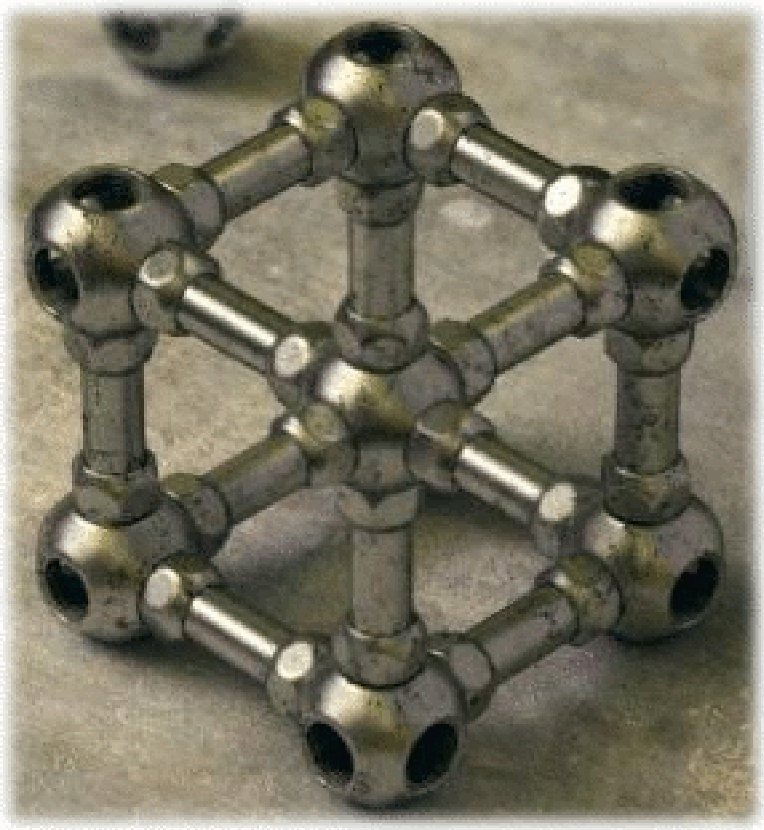

The Perception and Content of Cast Shadows: An Interdisciplinary Review [1]

Author(s):
Hannah M. Dee
Paulo E. Santos

[1]This work was supported by Fapesp Project LogProb, grant 2008/03995-5, São Paulo, Brazil.

Spatial Cognition & Computation

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hscc20

The Perception and Content of Cast Shadows: An Interdisciplinary Review

Hannah M. Dee [a] & Paulo E. Santos [b]

[a] Computer Science, Aberystwyth University, Penglais, Aberystwyth, UK

[b] IAAA, Centro Universitário da FEI, Electric Engineering Department, São Paulo, Brasil

Available online: 17 Aug 2011

To cite this article: Hannah M. Dee & Paulo E. Santos (2011): The Perception and Content of Cast Shadows: An Interdisciplinary Review, Spatial Cognition & Computation, 11:3, 226-253

To link to this article: http://dx.doi.org/10.1080/13875868.2011.565396

Spatial Cognition & Computation, 11:226–253, 2011

ISSN: 1387-5868 print/1542-7633 online
DOI: 10.1080/13875868.2011.565396

The Perception and Content of Cast Shadows: An Interdisciplinary Review

Hannah M. Dee[1] and Paulo E. Santos[2]

[1]Computer Science, Aberystwyth University, Penglais, Aberystwyth, UK
[2]IAAA, Electric Engineering Department, Centro Universitário da FEI, São Paulo, Brasil

Abstract: Recently, psychologists have turned their attention to the study of cast shadows and demonstrated that the human perceptual system values information from shadows very highly in the perception of spatial qualities, sometimes to the detriment of other cues. However with some notable and recent exceptions, computer vision systems treat cast shadows not as signal but as noise. This paper provides a concise yet comprehensive review of the literature on cast shadow perception from across the cognitive sciences, including the theoretical information available, the perception of shadows in human and machine vision, and the ways in which shadows can be used.

Keywords: shadows, perception, spatial reasoning, spatial perception

1. INTRODUCTION

Cast shadows are caused when a *caster* comes between a *light source* and a surface or *screen*. The information content in these types of shadows can therefore be used to provide knowledge about any or all of these three elements. As a very elementary example, if we assume that the light source does not move very fast and that the screen is flat and horizontal, we can draw conclusions about the size, motion and shape of casting objects by looking at their shadows. Casati (2004c) describes in depth the way shadows were used as powerful tools in early astronomical research for the determination of solstices and equinoxes, to provide an approximation of the distances from the Earth of the sun and moon, and to estimate the size and relative positions of celestial bodies. Keeping caster and screen constant, the motion of a light source has been used for thousands of years to measure time. Keeping the light source and screen constant, the use of shadows to inform about moving objects out of sight has been known for millennia—the allegory of the cave

Correspondence concerning this article should be addressed to Hannah M. Dee, Department of Computer Science, Llandinam Building, Aberystwyth University, Penglais, Aberystwyth, SY23 3DB, U.K. E-mail: hmd1@aber.ac.uk

from Plato (360 BC) concerns just this situation. In Galilean-era observations of the sky, shadows (and, in particular, eclipses) were used to show that the moon and the known planets were of the same nature as the Earth and that light has a finite speed and spreads by diffraction (as well as refraction and reflection). In the 20th century, shadows were used to verify the relativistic predictions of the deviation of light in the presence of mass and to suggest the hypothesis that the speed of the earth's rotation is slowing down.

In this paper we shall concentrate on the information content of cast shadows rather than self shading (where an object casts a shadow upon itself), and for the sake of brevity we shall refer to cast shadows as simply "shadows." These shadows are largely used by the human perceptual system to draw conclusions about everyday scenes, and as we shall see later in this paper (and according to recent psychophysical studies), some of these conclusions suggest that information from shadows can override conflicting depth cues present in the visual world. This implies that our perception of space is biased *towards* using information from shadows in certain situations.

In spite of this, computer vision systems have largely placed shadows in the position of noise to be filtered out.[1] In this paper we contrast the information available from shadows with state of the art computer vision methods for shadow filtering and shadow segmentation in order to make explicit the gap between what human perception deems as important to extract when constructing a spatial representation from a visual scene, and what current autonomous computer vision systems are designed to extract. Baxandall (1995) is an early interdisciplinary study, discussing the relationship between the representation of shadows during the Enlightenment (particularly within painting) and modern shadow perception, including the computational treatment of shadows. In contrast to the present work, however, the main concern of Baxandall is the discussion of the mid-18th-century thought on shadow perception and the technical literature on computer vision was searched only for specific issues; in addition, the field of computer vision has become much more involved in shadow perception in the intervening years and it is worth revisiting the question.

In order for the information content in shadows to be used as knowledge we note two difficult problems that a perceptual system has to solve first, which in turn give rise to a number of interesting questions that intersect cognitive science and computer vision research. The first problem is how shadows can be detected in the first place—some shadows have clear outlines and seem very "solid," yet we do not tend to misperceive shadows as objects. Other shadows have vague borders and therefore should be harder

[1]There is a large sub-field of vision research that deals with shape-from-shading whereby an object's self-shadowing is used to determine its shape, such as Kriegman and Belhumeur (1998). This line of research, however, does not take into account cast shadows. For more detail on shape from shading, see the recent review paper by Durou, Falcone, and Sagona (2008).

to perceive, but humans do not have any difficulty in doing so (as pointed out in Hering, 1878 and Bühler, 1922). Second, there has to be a consideration of the *Shadow Correspondence Problem* (Mamassian, 2004): given perceived objects and perceived shadows in one scene, how can shadows be unambiguously anchored to their casters? Once the detection and the shadow correspondence problems are solved we are left with the third major problem: what should we do with them? How are cast shadows used by the human visual system to determine the spatial characteristics of the scene? And how can machine vision systems exploit shadow information for spatial reasoning and analysis?

In this article we consider these three questions from an interdisciplinary perspective. Section 2 describes the various things we can learn from the investigation of shadows drawing upon optics and geometry. Section 3 moves on from the theoretical possibilities of shadow perception to consider evidence from the fields of art, computer graphics, psychology and neuroscience on the ways in which humans perceive shadows and the ways in which we actually use them. Section 4 considers the detection and use of shadows in computer vision, artificial intelligence and robotics, and finally Section 5 brings together the various interdisciplinary threads and provides pointers to open research questions.

2. THE INFORMATION CONTENT IN SHADOWS

Assuming that an environment has one strong light source (the primary light) and any other light sources are weak or diffuse (secondary light), the anatomy of a shadow cast upon a uniform screen is fairly simple: it consists of a main part (which is called the *umbra*), and a less dark fringe (the *penumbra*). The perceived darkness and any perceived color of the shadow depends upon the color of the screen, the intensity of the primary light source and the intensity and color of any secondary (or "ambient") illumination. The width of the penumbra depends upon the size of the primary light source, and the distance from the caster to the screen. A diagram of the shadow formation process is given in Figure 1. The situation becomes more complicated in the presence of multiple strong light sources, but similar principles apply. It is worth pointing out also that Figure 1 shows a simplification of the shadow formation process, since it ignores the effect of diffraction (which makes the shadow of an object slightly bigger than that provided by linear projection). This effect is minor, and can be considered to be irrelevant for the (human or machine) perception of shadows and for reasoning about shadows in the commonsense space, which are the main concerns of this work.

In real-world scenes a detailed model of shadow formation needs to take into account a number of different factors, related to the caster, light source and screen:

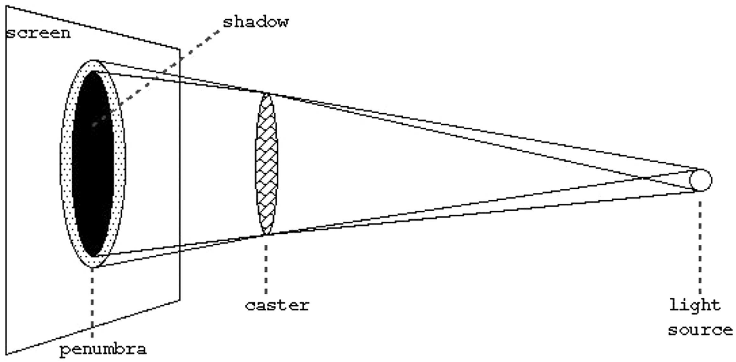

Figure 1. The anatomy of a shadow. The shadowed area is totally occluded from the primary light source by the caster, and the penumbra is partially occluded (that is, from the penumbra it is possible to see some part of the light source). Black lines indicate lines of sight. With a point light source, there is no penumbra.

- **Caster information:**
  - The shape and size of the caster determine size and shape of shadow;
  - The position (and pose) of the caster, particularly with respect to the light source, affects the shape, size and location of the shadow;
  - Opaque objects cast solid shadows, but translucent objects cast colored or weak shadows.
- **Light information:**
  - The shape and size of the light source determine characteristics of the penumbra;
  - The position of the source (along with the position of the caster) determines location of the shadow;
  - Light source intensity determines the contrast between shaded and non-shaded areas;
  - The intensity of any ambient illumination also affects contrast;
  - The color of ambient illumination determines the color of the shadow.
- **Screen information:**
  - Screen orientation with regard to light source determines the degree of distortion in shadow shape;
  - The shape and location of background clutter can cause shadows to split, distort, or merge.

By making assumptions about or keeping constant some of these factors, shadows can be used to determine various aspects of the visual scene. Casati (2004b) overviews the information encoded in shadows, which is not necessarily exploited by our perceptual system. For instance, the observation of a shadow in a scene, but not its caster, indicates the presence of objects outside the visual field (or occluded objects). Shadows indicate the direction

(a) (b)

Figure 2. In some situations shadows carry information about objects outside of view, via the "viewpoint" of the light source. In a), we can conclude that there is someone out of the scene behind the observer, and in b) we can conclude that there is an object "hiding" behind the pot. Photo 2(a) shows the artwork *Shadow Stone*, by Andy Goldsworthy, a work of art which encourages viewers to play with shadows (color figure available online).

in which the light source can be found, and intensity of the source (or the relative intensity of multiple sources). The width of the penumbra informs about the angular size of the source, and the distortion of the shadow outline (with respect to the shape of the caster) indicates the texture of the screen. Shadow motion carries information about the 3D structure of a caster, about the caster's motion in depth or about the geometry of the screen. Another important fact about the information content of shadows is that they can be seen as providing the observer with *a second viewpoint*: that of the light source, as the shadow depicts the projection of the caster's terminator line. This idea is illustrated in Figure 2.

3. THE HUMAN PERCEPTION OF SHADOWS

In this section we discuss the ways in which the human perceptual system handles shadows, considering evidence from art history, computer graphics, psychology and neuroscience. The question "*what makes a dark patch in a scene shadow-like?*" is not a simple one, and the human ability to make complicated judgements about 3D location in space based upon shadows has to be contrasted with our ability to perceive fundamentally inaccurate dark patches as shadows.

3.1. Depiction in Art

A glimpse of how the human perceptual system uses the information from (static) cast shadows can be obtained from the analysis of artistic depictions of the natural world. As Conway and Livingstone (2007) point out, in order to translate a convincing *impression* of the external world, artists explore rules of perspective, of color perception or visual illusions. Some of the information found in cast shadows was intensely explored by painters during the Renaissance, mainly in order to depict the position of important objects in scenes or to represent relative depth (Costa Kauffman, 1979). Indeed, Leonardo da Vinci himself carried out many observations into the way in which shadows are cast (for example, explaining why shadows cast by the sun on a white wall tend to look blue) and was also probably the first to relate the appearance of shadows with occlusion, when he says "no luminous body ever sees the shadows that it generates." The influence of da Vinci's work on shadows is discussed in detail in Fiorani (2008).

In particular, it was through the investigation of how the 3D world could be depicted in 2D paintings that projective geometry came to be developed in the 15th century, although Costa Kauffman (1979) argues that it is unclear whether the observation of shadows as projections played a central role in the development of this discipline.

It is worth mentioning the lack of cast shadow depiction in middle-age European art and in (pre-20th century) non-Western cultures.[2] Shadows were depicted in Hellenistic and Roman paintings, as seen in mosaics from Pompeii, but then largely disappear from the art historical record after the fall of the Roman empire. Once shadows reappear in the artistic world, a great number of shadow depictions are physically impossible (e.g., Figure 3). This shows us that once adopted by painters, the use of shadows was far from straightforward.

This neglect of shadows in artistic representations of the world could be explained by the inherent difficulty of depicting the right characteristics of luminosity (and imprecise borders) to make dark patches on canvas be perceived as shadows, as argued in Casati (2004c, 2006), or it may be due to the fact that the human perceptual system is simply insensitive to some of the information provided by static cast shadows.

Considering this point, Jacobson and Werner (2004) have investigated how sensitive our visual system is to static cast shadows using a visual search experiment in which human viewers had to determine which shadows were "impossible" in scenes with a number of casters and shadows. The results of this experiment indicate that the subjects were generally insensitive to inconsistencies in cast shadows, from which the authors concluded that the inclusion of cast shadows is not critical to the understanding of pictorial

[2] Although a very worthwhile mention here has to go to Chinese shadow puppetry. While this art, strictly speaking, is concerned with *using* and not *depicting* shadows, it has been around for millennia.

Figure 3. A cropped and contrast-enhanced portion of "*St. George killing the dragon*," Enea Vico, 1542. Note the shadow cast by the horse upon the ground; not only does it look *more* like the rear of a horse than the real shadow would, it is also an impossible shadow as St George's shadow is absent.

art. Cavanagh (2005) suggests that those transgressions of standard physics in visual art that pass unnoticed by the viewers' understanding (such as inconsistent shadows) indicate that our perceptual system uses a simplified physics to interpret the world. This simplified physics facilitates an efficient assessment of the visual world. Taking a different view, Casati (2007) argues that impossible shadows, often drawn as replicas of objects ("*copycat*" shadows, see for example Figure 3), are better cues for the localisation of casters in scene depictions than a more realistic shadow. This observation seems to contrast with Cavanagh's hypothesis of simplified physics, as the visual processing of replicas of objects corresponds to a more complex visual situation than that found in everyday life.[3]

3.2. Computer Graphics

Closely related to the painters' need to depict shadows, computer graphics is also interested in the rendering of the spatiotemporal structure of scenes, and

[3]That is, the physics necessary to cast a copycat shadow is richer than standard physics, and so cannot be a simplified physics as Cavanagh suggests.

therefore it has considered the determination and rendering of cast shadows in great depth. The first survey on shadow algorithms for computer graphics was presented in Crow (1977), which provides a classification of the early methods. A more up-to-date survey is presented by Woo, Poulin, & Fournier (1990), where shadow algorithms are classified by the type of shadows they produce: hard shadows, soft shadows, shadows of transparent objects and shadows for complex modelling primitives. In general, the large majority of shadow algorithms are based on the following methods: area subdivision, ray tracing, radiosity, shadow volumes and shadow maps (or z-buffers). Yet more recently, Hasenfratz, Lapierre, Holzschuch, and Sillion (2003) survey real-time "soft-shadow" algorithms. Rendering soft shadows realistically is a hard problem, and none of the modern algorithms cope with all of the difficulties involved in this task. Instead of trying to produce realistic shadows, Sattler, Sarlette, Mücken, and Klein (2005) evaluate the level of complexity required to produce shadows that are sufficiently detailed to be acceptable by the human perceptual system, with the final aim of using simplified models of scene objects to reduce the complexity of shadow rendering. Much computer graphics work provides shadows which seem realistic at first glance, but which become less so upon detailed inspection. For instance, much shadow rendering in 3D computer games is done by simple shadow maps whereby the effects of multiple reflectance in the environment is ignored. As a consequence, the shadowed regions look the same when observed from distinct angles (they should appear more diffuse the farther away they are from the observer). Additionally, crevices and depressions in objects are often either treated as dark patches, or as bright as the rest of the object.[4] These examples also suggest that the human perceptual system does not attend to every aspect of shadows, but (as the experimental evidence we are about to consider confirms) uses cast shadows to determine the 3D spatial organisation of a scene.

3.3. Experimental Studies

Several recent results from experimental psychology suggest that the human perceptual system prefers cues provided by shadows over other information in order to infer 3D motion of objects. Surprisingly, shadows are trusted *more* than changes in apparent object size. In one experiment presented by Kersten, Mamassian, and Knill (1994), a number of human subjects were presented with a computer simulation in which the shadow of a static square (cast on a chequered screen) moves away from its caster. Most subjects reported perceiving the square moving towards and away from the background according to the shadow motion, even though the size of the square remained unchanged throughout the experiment (this was clear from the static chequered background). It is worth pointing out that, geometrically, there are a

[4]http://gizmodo.com/5582218/what-directx-11-is-and-what-it-means-to-you

number of possible competing hypotheses for the shadow motion that would be more coherent than object motion as an explanation in this case (e.g., the motion of the light source). However, subjects even reported having the illusion of the object changing in size according to the shadow's motion. See Figure 4 for a static example of this visual effect.

Further situations were explored in Kersten, Knill, Mamassian, and Bülthoff (1996) to verify the effect of shadow perception on the perception of motion in depth. Subjects were shown two distinct animations of a ball moving inside of a box. In the first animation, the ball was made to move along a diagonal inside the box, while the ball's shadow described an horizontal trajectory in the image. In the second situation, the ball's trajectory was the same, but the ball's *shadow* moved in such a way that it was aways connected to its caster. Even though the ball's trajectory was identical in both situations, and there was no change in the size of any objects in the scene during motion, all observers interpreted the ball as rising above the floor in the situation where the shadow motion was horizontal, but as receding in depth in the other. These findings (summarized in Mamassian, Knill, & Kersten, 1998) suggest that, in some cases at least, the human perceptual system is biased to use shadow information for the interpretation of 3D motion and that shadow information can even override notions of conservation of object size.

As well as providing a strong cue about motion in depth, cast shadows provide information that could be used in the interpretation of surface shape of the screen, however the experimental findings of Kersten, Mamassian and colleagues suggest that this information is not used by the human perceptual system.

Psychological studies investigating the relationship between shadow perception and object recognition tell a less clear story. Braje, Legge, and Kersten (2000) report results from three experiments involving the recognition of natural objects with shadows in several experimental conditions, and suggest that human object recognition is not affected by the presence of shadows. The authors conclude that the results are consistent with a feature-based

Figure 4. Still image version of the experiment presented in Kersten et al. (1994), in which only the shadow changes but we perceive the square as moving.

representation of objects, where shadows may be filtered out as noise. However, it may also be the case that the results obtained are dependent on the type of stimuli used in the experiments (simple familiar objects), which contained much redundant information that could reduce the importance of the information provided by object's shadows. Castiello (2001) reports an experiment with contrasting results, in which the perception of objects is hindered when presented with incongruent cast shadows (wrong shadow) or incongruent lighting with respect to the shadows (shadow on the wrong side with respect to the light source). There are two competing explanations for these findings: either the perception of shadows is used to improve object recognition in certain situations (and, therefore, adding an extra-level of processing to perception), or incongruent shadows work as distractors in the scene.

Another area of research considered by psychologists is that of the spatial relation between shadow and caster, particularly concerning the determination of optical contact.[5] In particular, Ni, Braunstein, and Andersen (2004) investigate the difference in depth perceptions of a floating object with relation with an object on the ground following it "like a shadow." The authors want to address three fundamental questions: What are the features that make a shadow be perceived as such? What is the effect of object separation in the perception of depth from shadows? In situations with multiple shadows, what are the features that make us associate one particular shadow with an object? They investigate these questions by varying the light intensity of the lower object, its thickness and its motion relative to the casting object. Perhaps unsurprisingly, darker objects are more readily perceived as shadows. Common motion of object and shadow is an important feature for shadow association and, in situations with multiple shadows, the authors suggest that common speeds decide which shadow is associated with an object (relative size being a secondary concern).

Psychological research also suggests that our perceptual system uses cast shadows as a coarse cue: it does not matter if the shadow is the wrong shape for the casting object, it just has to be associated with the caster, telling a coherent story about the object motion or location. Enns and Rensink (1990) were the first to investigate the effect of unusual or "wrong" shadows on our perception, by creating images in which one object had a shadow which was inconsistent with the types of shadows we see day-to-day. Bonfiglioli, Pavani, and Castiello (2004) carried out a naturalistic study using real objects with fake shadows, and discovered that shadows do not affect our verbal reports of what is going on, but can affect the way we reach for an object; shadows which are the "wrong" shape affect our physical behavior but not our verbal reaction times. Ostrovsky, Cavanagh, and Sinha (2005) also investigate shadows which arise from inconsistent illumination. These studies involve

[5]Optical contact is the place where an object is connected to the background in a 2D projection of a 3D scene.

the presentation of an array of identical objects with consistent shadows and shading, but with one drawn as though it is lit from a different direction. While the earlier studies suggested that the illumination change was easy to detect (indeed, it *popped out*) the later study shows this may have been an artifact of the regularity of the array. When objects are arranged in a more random fashion, the shadow difference is harder to perceive (as long as the shadow is plausible). More recently, Farid and Bravo (2010) consider the human ability to detect images that have been digitally manipulated, and present evidence that the human perceptual system is not capable of detecting simple inconsistencies on the position of light source, caster and shadow. Interestingly, Casati (2006) comes to a similar conclusion through the observation that dark patches in paintings, sometimes bearing no resemblance to real shadows, suffice to enhance the perception of depth. The visual system seems to extract a position estimation from shadows early on in processing, then filters them out in order to avoid interpreting shadows as objects in further spatial inferences.[6]

Rensink and Cavanagh (2004) present compelling evidence for the hypotheses that shadows are processed early in the visual pathway and then discarded, and that we use an assumption of a single overhead light source in doing this. Using a visual search methodology, they show that the detection of shadow-like shapes consistent with an overhead light source takes longer than the detection of the exact same shape in other situations. If the shape is altered so it is not shadow-like (it is lighter, or has the wrong texture, or the wrong edge features to be a shadow) or the shape is shadow-like but is consistent with illumination from below, visual search is much quicker. This is consistent with the hypothesis that regions not recognized as shadows were still available for rapid search, whereas shadow-like regions were discounted early in visual processing and thus had to be processed consciously to accomplish the visual search task. Therefore, it seems that shadow processing is both implicit (i.e., without conscious awareness) and automatic (i.e., without attention): observers cannot stop interpreting appropriate regions as shadows, even when this gets in the way of using information in the image.

The question of whether shadow processing is implicit or not is considered in Castiello, Lusher, Burton, and Disler (2003) in which, through analysing cast shadow perception in groups of people with brain injuries, the authors try to localise cast shadow processing in the brain and to determine whether conscious awareness is necessary. They show that the performance in a simple object recognition task is hindered if the shadow is missing or incongruent (does not match the object). This effect exists even in brain-injured patients suffering from visual neglect, who are not aware of the existence of the shadow. These findings suggest that our ability to process

[6]It is worth mentioning that even if shadows are discounted, there is no evidence that this discounting may affect functions other than object identification. The information in shadows could (at least in principle) still be used for depth estimation.

and deal with cast shadows is not dependent upon our conscious awareness of them and, therefore, is an implicit process. Furthermore, the authors test the hypothesis that shadow processing in the human brain is located in the temporal lobe (following some previous evidence that an analogous process occurs in monkeys' temporal areas). For that, a number of patients suffering from left visual field neglect caused by lesions in the temporal lobe are subject to the same object recognition task (with incongruent shadows) as patients with frontal lobe lesions. In this case, the temporal lobe patients had lower reaction times when presented with shadows to the left of an object, providing some support for the hypothesis.

Whether shadow processing is implicit or explicit, there is evidence that shadows cast by a person's own body parts are used more effectively in judgements about extra-personal space than shadows from other objects carrying analogous information. Evidence for ths comes from Pavani and Castiello (2004), in which the judgement of distances from shadows of the subject's own hand diverged from similar judgements when the subjects were wearing a polygonal glove. Following a similar experimental setup, Galfano and Pavani (2005) find support for the hypothesis that body-shadows act as cues for attention.

4. THE MACHINE PERCEPTION OF SHADOWS

In this section we provide an overview of the main algorithms and research topics within computer vision for shadow detection. In this, we place more emphasis upon those systems which use shadows as information than those aimed at filtering shadows as noise. We also discuss some literature in artificial intelligence and robotics in which cast shadows are considered.

The first paper to attempt a formalisation of shadows from a machine vision standpoint is that of Waltz (1975). This paper presents a number of computer programs capable of reconstructing 3D descriptions from line drawings of objects and their shadows. After an initial identification and grouping of shadow lines and regions from line drawings, the proposed system is capable of extracting high-level relations representing contact, support and orientation between objects.

Much shadow detection work in computer vision, however, is centred around the idea of shadow as noise. Two broad approaches are affected by shadows: the first deals mainly with single images and is associated with the segmentation of images into the objects that they depict; and the second deals with video and is concerned with the identification of moving objects. Shadows are problematic in both cases—they cause spurious segmentations in the first instance, and spurious foreground objects in the second. Perhaps the simplest shadow detection method proposed is that of Troccoli and Allen (2004), in which a grey-scale image is simply thresholded and the darker pixels are labelled "shadow." In the archaeological images the authors deal with, this works reasonably well; however, for more complex images more sophisticated algorithms are called for.

Those algorithms dealing with single images use color and texture information to group image pixels into regions that correspond to single elements in the real world (such as grass, or trees). This can be seen as an exercise in *color constancy*; the aim is to determine the color of the underlying object in various light conditions, and in this context shadows are merely one of these light conditions rather than an object of study in themselves. The existence of strong shadows can cause spurious segmentations, and so shadow detection is performed in order to classify shaded pixels as part of the *screen*, rather than as shadow. An example of this sort of work is that of Vazquez, Weijer, and Baldrich (2008) who engage not so much in shadow detection as in shadow *blindness*. The aim is a segmentation in which image components are classified regardless of self-shading and inter-shading; this is achieved by identifying "ridges" in color space that are characteristic of a particular dominant color under differing lighting conditions. While these ridges could conceivably be used as part of a shadow detection algorithm, this is not part of their current work. We propose that shadow removal algorithms such as those introduced by Finlayson and colleagues (e.g., Finlayson, Hordley, Drew, & Lu, 2006; Finlayson, Fredembach, & Drew, 2007) fall in a similar category—they are concerned with shadow blindness, and only work on individual images (and are often too slow to be considered useful for video processing, or use "tricks" such as photographing the same scene twice with different colored filters).

The second major consideration of shadows within computer vision comes when detecting moving objects. This is commonly done by subtracting "background" from video to find objects of interest, where background is detected by finding those pixels or image regions, which do not change much in color. In doing this, shadows become a major source of false positives as a cast shadow will make an otherwise uninteresting pixel change color.

Thus in this sub-field of computer vision, shadow detection almost always involves some model of the color of the *screen*, or *background*, and then detection is performed using a model of shadows characterizing them as "roughly the same color as background, but darker." Prati, Mikic, Trivedi, and Cucchiara (2003) provide an overview and a taxonomy of shadow detection techniques, dividing them into *model-based* and *non-model-based* and then further into *parametric* and *non-parametric* techniques. This categorization does not apply so well to more recent works, many of which can be thought of as "ensemble methods." Thus we make a different distinction, between methods which detect shadows based upon color information alone, and those which incorporate some form of spatial or spatiotemporal information (such as the relationship between pixels classified as shadow, or the spatial relationship between known objects and shadow regions). As we have seen in Section 3, the human visual system uses not only color but also texture, motion, and spatial organization when dealing with shadows.

Cucchiara, Grana, Neri, Piccardi, and Prati (2001) take as their starting point detected moving objects and a background model. The pixel values of moving objects are converted to the HSV (Hue, Saturation, and Value) color space, and then observed values of all three HSV components are compared to those of the background model. The particular calculations they make are the difference between foreground and background values for H and S, and the ratio of the two V values. This captures the intuitive observations that shadows are about the same hue as the same part of the scene unshadowed, slightly more saturated, and darker. Stauder, Mech, and Ostermann (1999) use assumptions about the background (it will dominate the scene), the nature of shadows and luminance (shadows are darker and tend to have uniform shading) and the presence of moving and static edges. In addition to these they use the width of edges to detect a shadow's *penumbra*: in a world without point light sources, shadows have fuzzy edges—so those regions bound by broad edges are candidates for shadows as the edges could be penumbra.

Martel-Brisson and Zaccarin (2007) present a Gaussian mixture model based approach for shadow detection. They use three types of model in coordination to find the shadows: one of these represents the physical characteristics of shadows, and the other two capture statistical properties of the way colors are expected to vary when shaded and unshaded. The simplest is a physical model of shadow appearance, which essentially expresses the familiar notion that shadows are similarly colored to background but darker, and for this they make use of earlier techniques (e.g., Cucchiara et al., 2001; Hoprasert, Harwood, & Davis, 1999). This alone is insufficient, and they augment the physical representation with statistical learning to try to minimize false shadows. Using a Gaussian mixture model (GMM) with four Gaussians to model the distribution of pixel colors in the background, they assume the most stable component is the actual background and all others foreground. As observations accrue, various other colors will be captured by the GMM as occurring at this one particular pixel. However, the shadowed value can be assumed to be the most stable foreground Gaussian as it will occur more frequently than any foreground color caused by moving objects or noise. This most stable foreground component is then compared to the physical shadow model, and if it is a plausible shadow color, the learning parameter of that particular Gaussian is increased so that distributions which are plausible shadow colors at a particular pixel converge more quickly. Their third component (the Gaussian Mixture Shadow Model, or GMSM) stores the parameters of up to three previously learnt stable shadow Gaussians, which avoids the "forgetting" of shadow characteristics in periods of great foreground motion or changing illumination.

Joshi and Papanikolopoulos (2008a) present work which uses a support vector machine (SVM) to perform classification of image regions into shadow and non-shadow categories. As with many of the papers we discuss here, their starting point is a GMM of background appearance and a "weak classifier." Their classifier is based upon color and edge features, and is used to train the

SVM. This allows for more variation in shadow appearance than many other approaches, as an SVM can learn a more complicated discriminatory function. In Joshi and Papanikolopoulos (2008b) this approach is extended using a co-training framework. In co-training, a small set of labelled examples are used to train a pair of classifiers, and then for previously unseen and unlabelled examples the output of each classifier is used as new labelled data to train the other. The two classifiers presented in this work use edge features and color features, and thus those patches which are confidently classified as shadow based upon color are used as new examples for training the classifier based upon edge features, and vice versa. The presented results are very impressive.

Physics-based techniques and features for shadow modelling have become more popular in the last two years. Martel-Brisson and Zaccarin (2008) take a simplified reflectance model and use it to learn the way in which colors change when shaded, and Huang and Chen (2009) have also incorporated a richer, physics-based color model for shadow detection based upon the work of Maxwell, Friedhoff, and Smith (2008). Maxwell et al. present a bi-illuminant dichromatic reflection model, which enables the separation of the effects of lighting (direct and ambient) from the effects of surface reflectance. Huang and Chen simplify this model in several ways, such as assuming that the ambient illumination is constant, which enables them to implement shadow detection based upon the simplified model in a video analysis task. Their system involves a global shadow model, which is a GMM representing the change in color of a pixel when shaded (based upon the ambient illumination), and a per-pixel color model. The use of a global model means that their approach is very fast to train and robust to low frame rate videos. Results are presented which show that this approach performs comparably to other methods (including Martel-Brisson & Zaccarin, 2008; Liu, Huang, Tan, & Wang, 2007; Martel-Brisson & Zaccarin, 2007).

4.1. Using Spatial Information for Shadow Detection

We now move on to techniques which incorporate spatial information. The simplest way to do this is to use some measure of "spatial coherence" (shadowed pixels tend to be next to other shadowed pixels), but some authors use more sophisticated spatial models of shadow location including assumptions about the light location or the relationship between shadow and caster.

Porikli and Thornton (2005) present a method which is similar in spirit to that of Martel-Brisson and Zaccarin's earlier work (2005, 2007). They also use a physical model of shadows as a weak classifier (shadows are darker than the expected background), and use those pixels which satisfy this condition for updating their Gaussian shadow models. However they introduce a spatial coherence condition in addition to color information, capturing the basic idea that shaded pixels are more likely to be found next to other shaded pixels.

Nadimi and Bhanu (2004) describe a technique for shadow detection which is partly statistical and partly based upon physical attributes is described. This is a seven-stage algorithm which is novel within computer vision as it models the physical characteristics of shadows from two light sources: "diffuse" and "point" (the sky and the sun, respectively). They start with a GMM-based moving object detector, then reduce the detected pixels by getting rid of those which are brighter than the corresponding background pixel. Next, the detected area is thresholded on saturation, and if not too saturated they keep pixels which are bluer (shadows are assumed to be illuminated only by sky, not sun). They then use a new "spatio-temporal albedo" measure that looks at neighboring pixels in time and space, searching for those which are uniform. Remaining pixels are candidate shadow pixels, and the difference between these and background pixels is used to discard those which are actually background. The penultimate step estimates body color from a segmented region, and the final step matches body color against learnt body colors from the scene. This technique seems to work well on the author's test data, but is limited to outdoor situations.

Mikic, Cosman, Kogut, and Trivedi (2000) also use spatial coherence. This is enforced by smoothing and morphological operations, to eliminate small shadow regions that occur inside foreground or background. Their color based classifier is founded upon the observation that the color change due to shading can be approximated by a diagonal matrix transformation in color space. Rittscher, Kato, Joga, and Blake (2000) enforce spatial coherence through the use of a Markov Random Field (MRF); they also use temporal continuity constraints in their shadow and foreground detection. Salvador, Cavallaro, and Ebrahimi (2004) also exploit spatial coherence. Shadow pixels are initially detected based upon color difference to a reference pixel.[7] They use an observation window rather than working at the level of the individual pixel to reduce noise, and a Gaussian distribution to model the difference between shadow and non-shadow pixel colors. They then use spatial constraints to remove spurious object pixels classified as shadow (e.g., shadow regions cannot be entirely surrounded by object regions[8]), and a final information integration stage makes the decision as to whether a pixel depicts a shadow or not.

A similar effect is obtained in Wang, Loe, Tan, and Wu (2005) and extended to incorporate edge information in Wang, Loe, and Wu (2006). This work uses a statistical approach based upon both Hidden Markov Models and Markov Random Fields. They combine these two models in a Dynamic Hidden Markov Random Field (DHMRF). The dynamic segmentation is modelled within the Hidden Markov Model framework, and spatial constraints are handled by the Markov Random Field. This has the effect of making a

[7]In video, the reference pixel is at the same spatial location but from the background model, in a still image, the reference pixel is a neighbor.

[8]This rule is not true in all cases; we must assume that the authors were not considering objects with holes.

pixel more likely to be shadow if its neighbors are shaded. They model background variation using a GMM; when a foreground pixel is discovered they use the DHMRF framework (based upon color, spatial coherence and also edges) to decide whether that pixel is background (and hence update the GMM) or whether it is shadow or foreground. A further enhancement is introduced in Wang and Ye (2008) in which additional latent variables are introduced further stabilising the segmentations. As the authors are considering road traffic scenes alone they impose a further constraint by assuming that foreground objects are rectangular. Similarly, Liu et al. (2007) use GMMs and weak spatial information, with the spatial information encoded using an MRF to smooth detected shadow pixels. Benedek and Szirányi (2008) incorporate texture within their framework using kernel methods (rather than edge methods), and also use an MRF formulation to perform smoothing.

Hsieh, Hu, Chang, and Chen (2003) propose a technique that uses a stronger form of spatial information (as well as color information). They assume that the object casting the shadow is a pedestrian and that the shadow is being cast onto a flat planar surface (that is to say, the ground). They first perform a background subtraction then morphological operations to obtain the moving people and their shadows. On this segmented area they then calculate the center of gravity and orientation using moments. This allows them to find a rough segmentation of person from shadow by finding the bottom of the person and drawing a diagonal line (oriented to match the orientation of the entire segmented area): see Figure 5(a) for an illustration of this. Given this rough shadow segmentation they then build a Gaussian model of the color distribution of the shadow pixels, allowing color based refinement of the shadow model.

Renno, Orwell, Thirde, and Jones (2004) describe a shadow detection technique which uses strong spatial information to augment a color based shadow segmentation. In this article they deal with the characteristic quadruple shadows cast by football players under floodlights, and a novel skeletonization approach is used to distinguish those foreground detections due to the cast shadows (which appear on the floor) and those which are due to actual foreground motion. Those pixels which are *most likely* to be shadow are used to train the shadow GMM, and the others to train the foreground models. Figure 5(b) illustrates this approach.

Both Hsieh et al. (2003) and Renno et al. (2004) use strong spatial information, but also make some strong assumptions about the light, the screen, and the caster. Hsieh et al. have difficulty in detecting shadows where there are overlapping pedestrians, or pedestrians assuming unusual poses (sticking their arms out, for example), and explicitly only model shadows cast on a planar surface by people. Renno et al. make similar assumptions—given their domain (soccer tracking) this is a reasonable thing to do as soccer players are usually human and soccer pitches are planar with a characteristic lighting pattern. In scenes in which these assumptions do not hold, these approaches will naturally have difficulty.

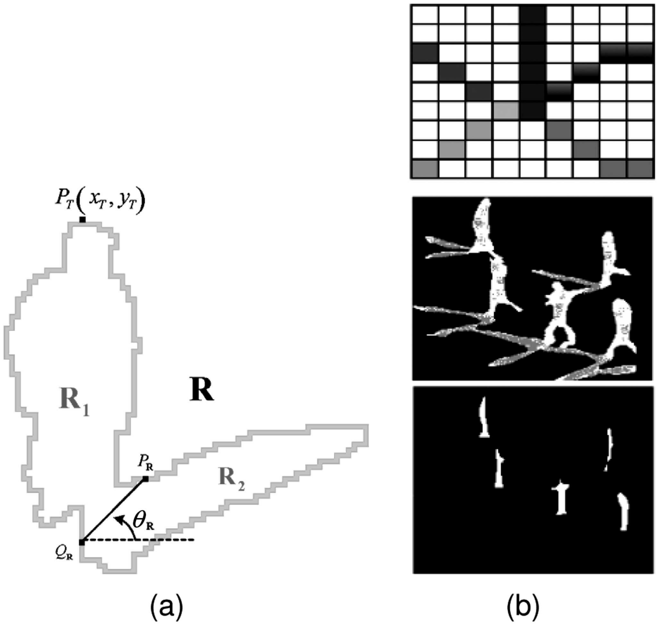

Figure 5. Two computer vision techniques that exploit *strong* spatial as well as color information to perform shadow removal: 5(a) shows the method of Hsieh et al. (2003), and 5(b) shows the skeletonisation (top) intermediate segmentation (middle) and final results of shadow removal (bottom) from Renno et al. (2004b).

Cucchiara, Piccardi, and Prati (2003) describe an extension to their earlier work (Cucchiara et al., 2001) in which a higher level reasoning component classifies regions as one of *Moving object*, *Background*, *Shadow*, *Ghost* or *Ghost shadow* by incorporating spatial constraints upon the arrangement of regions within a higher-level reasoning component. Regions classified as shadow have to be adjacent to moving object regions. This prevents spurious shadows unattached to casters being "invented" by the software. In their terminology, a ghost is an artifact of the tracking system and can correspond to an erroneous foreground detection or a an erroneous shadow detection. By using the reasoning component to work out where shadows and ghosts *should* appear, they handle these problems well.

When we consider systems which *use* shadows, instead of filtering them out, there are only a handful: Cao and Foroosh (2007) use known 3D locations and their cast shadows to perform camera calibration and light location (using known casters and screen to tell about light source); Caspi and Werman (2006) use the moving shadows cast by known vertical objects (flagpoles, the sides of buildings) to determine the 3D shape of objects on the ground (using the shadow to tell about the shape of the screen).

Balan, Black, Haussecker, and Sigal (2007) use shadows as a source of information for detailed human pose recognition: they show that using a single shadow from a fixed light source can provide disambiguation in a similar way to using additional cameras. They estimate human pose from a single calibrated camera, using a strong light source to cast shadows on the ground. The shape of the shadow and the shape of the observed silhouette taken together enable detailed recovery of the pose of the human. In this work they also discuss the estimation of light source position (given pose and shape), and the surface reflectance of the person under consideration.

4.2. Shadows in Robotics

Within robotics, the emphasis of the computer vision task shifts from the passive interpretation of a scene to active exploration of the visual world and the robot's place within it. Perhaps unsurprisingly the *use* of shadows within robotics is therefore more common than within mainstream computer vision. There are several systems which make use of cast shadows for informing about the location of the robot or the robot's manipulators, and the relationship between the robot and its environment.

Two systems have used the shadow cast by a robot's arm to refine the robot's estimation of limb location. When a robot wishes to move its arm from A to B in the real world, it has various sources of information about the motion. Visual feedback is a central part of this and these recent papers have incorporated shadows into the visual element of robot motion control, inspired in part by Castiello et al. (2003), who showed that humans use the shadows of their own limbs in a similar fashion. Fitzpatrick and Torres-Jara (2004) track the position of a robotic arm and its shadow cast on a table to derive an estimate of the time of contact between the arm and the table. Shadows are detected in this work using a combination of two methods: in the first, a background model of the workspace is built without the arm and then used to determine light changes when the arm is within the camera view. The second method compares subsequent frames in order to detect moving regions of light change. The authors motivate their work pointing out that depth from shadows and stereopsis may work as complementary cues for robot perception, while the latter is limited to surfaces rich in textures, the former works well in smooth (or even reflective) surfaces. Cheah, Liu, and Slotine (2006) present a novel controller for a robot manipulator, providing a solution to the problem of trajectory control in the presence of kinematic and dynamic uncertainty. To evaluate their results, an industrial robot arm was controlled using the visual observation of the trajectory of its own shadow.

In a similar vein, Kunii and Goton (2003) propose a *Shadow Range Finder* system that uses the shadow cast by a robot arm on the surface of a terrain in order to obtain depth information around target objects. In planetary explorations this type of system may provide low-cost, energy-saving sensors for the analysis of the terrain surrounding rock samples of interest.

Within the field of robotics planning and navigation, Tompkins, Stentz, & Whittaker (2001) describe an autonomous path planning system that takes into account various conditions of the robot's state, including peculiarities of the terrain and lighting. In this context, the information about shadows cast by terrain irregularities allows the rover to plan a trajectory that maximises the trade-off between the exposure of the solar cells to sun light and the limited resources (including time) in planetary missions. More recently, Santos, Dee, and Fenelon (2009) describe an initial representation of cast shadows in terms of a spatial logic formalising occlusion relations. This initial representation is used in a mobile robot self-localisation procedure in office-like environments to determine the relative locations of light source, caster, and robot. In the context of industrial robotics, Lee, Roh, Kim, Moon, and Choi (2009) use cast shadows inside pipes to detect landmarks: by fitting bright lights to the front of their pipe inspection robot, they can determine when a pipe bends by detecting cast shadows.

5. CONCLUDING REMARKS

In this section we try to draw together the various approaches to shadow perception covered in this article, and suggest ways in which an interdisciplinary approach could guide future research. In particular, we consider the possible utility of holding evidence from human perception in mind when designing computer vision systems: might the short-cuts taken by our perceptual system provide clues for those researching artificial intelligence?

To summarize the psychological and neurological evidence discussed in Section 3, it appears that our visual system handles cast shadows by rapid processes in early vision[9] that extract coarse indicators of depth and 3D position in space, and then discards shadows just after so that they do not interfere in further processing (Rensink & Cavanagh, 2004). This is consistent with findings suggesting that shadow processing is implicit (Castiello et al., 2003), and with those results indicating that the human perceptual system does not rely on cast shadows for object recognition (Braje et al., 2000), even though this kind of information could (in principle) be used.

This rapid, pre-conscious processing of shadows is also in line with the apparent deficiencies in the early artistic depiction of shadows (Casati, 2004a), which could then be interpreted as the unavailability of shadow information during the conscious depiction of the 3D world on a 2D screen. This hypothesis is also in agreement with the existence of inconsistent or copycat shadows in paintings where these inconsistencies are imperceptible

[9]That is, in the first few hundred milliseconds of processing that does not involve stimulus-specific knowledge.

by the observers (Jacobson & Werner, 2004; Cavanagh, 2005; Casati, 2007). In other words, if the human perceptual system only extracts from shadows a coarse indication of 3D position in space very early in the processing pathway, and if the inconsistencies in the shadows are such that they can still be perceived as shadows, the coarse cues would be processed, and the inconsistencies would be discarded. The early vision processes only perceive a stimulus as shadow if it is a fairly homogeneous region, darker than the background, without internal edge features. Despite these constraints, we are still able to handle major variations in the appearance of shadows, perceiving as shadows those stimuli arising from inconsistent illumination or shape, or even thick dark patches in scenes (Ni et al., 2004; Elder, Trithart, Pintilie, & MacLean, 2004). Nevertheless, depth cues provided by shadows seem to have priority over other cues, such as the change in apparent object size during motion in depth (Kersten et al., 1994; Mamassian et al., 1998; Kersten et al., 1996).

The robustness to variation and rapid processing of shadows lead us to suggest that the interpretation of shadows as depth cues was incorporated into the human perceptual system at a very early stage of evolution. This hypothesis is also supported by recent research on animal cognition, whose results suggest that chimpanzees perceive depth using shadow information (Imura & Tomonaga, 2003, 2009). We conjecture that this is due to the need for rapid processing: perceiving every single aspect of a scene is a much harder procedure than just processing a coarse position estimate given by shadows. It is worth recalling here the evidence for an increase in reaction time of subjects when *fake* shadows interfere in object recognition (Castiello, 2001).

Saving processing by prioritizing shadows as depth cues is one idea from the human perception of shadows that could be used to enhance computer vision systems that have to deal with everyday scenes.

We know of no work to date within artificial intelligence or computer vision that uses shadows in the same way that human systems do. Within computer vision we can now find shadow detection algorithms using similar visual features to the human perceptual system (color and edge based features) and some spatial features (e.g., spatial coherence). However it remains the case that the aim of the vast majority of these systems is the ability to *ignore* shadows, not to *use* them. While the "*grand aim*" of computer vision can be stated as the semantic interpretation of images, the majority of current computer vision systems deal with sub-problems such as object recognition or motion detection. By filtering shadows out before extracting useful information from them, these systems could be losing important cues about the location of objects to be recognized, and the location of object motion within the scene.

This could be seen as a side effect of a wider trend within computer vision and artificial intelligence: while early systems drew on psychophysical results, in recent decades these fields have moved away from cognitively inspired design towards methodologies more aligned with statistics or engineering. Without wishing to take sides in the broader debate upon the foundations of AI, our thesis is a pragmatic one. *Some of the shortcuts taken by the human visual system are useful things to consider implementing within a computer vision context particularly when it comes to the perception of space and spatial relations*. In the case of shadows, as discussed in this paper, some noteworthy examples of this *appropriation* of human perception strategies are present in robotic vision (Section 4.2). However, the full information content of cast shadows (cf. Section 2) is yet to be used.

Ideally, we would envisage a system that rapidly identifies shadows in the scene, extracting implicit information so that this can be included in the representation of shadow casters for use in determining spatial relations and relative motion (the things for which humans seem to use shadows). The shadows could then be discarded, so that they would not interfere in further object recognition, as is usually done in computer vision (cf. Section 4). This idea requires the following components: a rough and rapid shadow identification method; a solution for the shadow correspondence problem; and a means of reasoning about spatial relations given the location shadows and shadow-caster correspondences.

The first of these components (the computational identification of shadows) is not trivial. However as Section 4 shows, much progress has been made. Techniques based on machine learning for obtaining a model of shadow characteristics can identify cast shadows in many situations. There are several ways in which these systems might be enhanced in light of the short-cuts that humans take—for example, by assuming an overhead light source, or by exploiting common motion of shadow and caster.

The shadow correspondence problem, however, looms over any attempt to incorporate notions of the relation between caster and shadow. This problem is non-trivial for several reasons: there may be various competing possible matches between shadows and objects in a complex scene (i.e., the shadow correspondence problem is underconstrained); the screen may not be planar, which may turn a point-to-point matching into a complex nonlinear registration procedure; and shadows of nearby objects may merge. A robust computational solution for the shadow correspondence is still an open problem.

One further piece of evidence in favor of computer vision scientists incorporating a model of shadow which includes the shadow's dependency upon its caster is that those systems that do incorporate a known caster (e.g., Renno et al., 2004) or which assume certain properties of the caster (e.g., Wang & Ye, 2008) perform very well indeed. Indeed one of the main conclusions we can draw from the psychological evidence is that human shadow perception is far from a linear process—shadow location and motion affects our perception of caster location and motion, but the converse is also true. It is therefore unsurprising that the machine perception of shadows is aided when some consideration of the caster is incorporated.

Finally, the third component of our ideal shadow system would be able to use the output of the first two components in order to reason and draw conclusions about the world from the information thus extracted. Having considered the psychophysical aspects of shadow perception, it is worth noting that computer vision systems are not necessarily limited by the way in which humans use shadows. Indeed they may also have much to gain by taking into account the entire information content in cast shadows (including that which is apparently unused by the human perceptual system). For instance, shadows can inform about the shape, size and pose of the caster; the position, intensity and shape of the light source; and the physical characteristics of the screen, given the distortion of the shadows cast on them. Some vision work does use shadows in this way (e.g., Balan et al., 2007) but this, as noted, is rare.

Given the value of computational theories of vision, and the way in which such theories can unify and connect interdisciplinary work, a challenge for future research is the combination of the various aspects of shadow perception into a general computational theory, along the lines of Marr's "*Vision*" (Marr, 1982). We hope this paper goes some way to starting this process.

ACKNOWLEDGMENTS

We would like to thank Roberto Casati, Ronald Rensink and an anonymous reviewer for their invaluable comments on an earlier version of this paper. We would also like to thank the British Council for funding exchange visits between the authors during which these ideas were developed.

Hannah Dee was partially supported by EPSRC Grant LAVID, EP/D061334/1, UK. Paulo Santos acknowledges support from FAPESP, LogProb 08/03995-5, and CNPq, Brazil.

REFERENCES

Balan, A. O., Black, M. J., Haussecker, H., & Sigal, L. (2007). Shining a light on human pose: On shadows, shading and the estimation of pose and shape. In *IEEE Proc. International Conference on Computer Vision (ICCV).* Rio de Janeiro, Brazil.

Baxandall, M. (1995). *Shadows and enlightenment.* New Haven & London: Yale University Press.

Benedek, C., & Szirányi, T. (2008). Bayesian foreground and shadow detection in uncertain frame rate surveillance videos. *IEEE Transactions on Image Processing, 17*(4), 608–621.

Bonfiglioli, C., Pavani, F., & Castiello, U. (2004). Differential effects of cast shadows on perception and action. *Perception, 33*(11), 1291–1304.

Braje, W. L., Legge, G. E., & Kersten, D. (2000). Invariant recognition of natural objects in the presence of shadows. *Perception, 29,* 383–398.

Bühler, K. (1922). *Die erscheinungsweisen der farben.* Jena: Fischer.

Cao, X., & Foroosh, H. (2007). Camera calibration and light source orientation from solar shadows. *Computer Vision and Image Understanding (CVIU), 105*(1), 60–72.

Casati, R. (2004a). Methodological issues in the study of the depiction of cast shadows: A case study in the relationships between art and cognition. *Journal of Aesthetics and Art Criticism, 62*(2), 163–174.

Casati, R. (2004b). The shadow knows: a primer on the informational structure of cast shadows. *Perception, 33*(11), 1385–1396.

Casati, R. (2004c). *Shadows: Unlocking their secrets, from plato to our time.* New York: Vintage Books.

Casati, R. (2006). The cognitive science of holes and cast shadows. *Trends in Cognitive Science, 10*(2), 54–55.

Casati, R. (2007). The copycat solution to the shadow correspondence problem. *Perception, 37*(4), 495–503.

Caspi, Y., & Werman, M. (2006). Vertical parallax from moving shadows. In *IEEE Proc. Computer Vision and Pattern Recognition (CVPR)* (pp. 2309–2315). New York, New York.

Castiello, U. (2001). Implicit processing of shadows. *Vision Research, 41,* 2305–2309.

Castiello, U., Lusher, D., Burton, C., & Disler, P. (2003). Shadows in the brain. *Journal of Cognitive Neuroscience, 15*(6), 862–872.

Cavanagh, P. (2005). The artist as neuroscientist. *Nature, 434,* 301–307.

Cheah, C.-C., Liu, C., & Slotine, J.-J. E. (2006). Adaptive tracking control for robots with unknown kinematic and dynamic properties. *I. J. Robotic Res., 25*(3), 283–296.

Conway, B. R., & Livingstone, M. S. (2007). Perspectives on science and art. *Current Opinion in Neurobiology, 17,* 1–7.

Costa Kauffmann, T. da. (1979). The perspective of shadows: The history of the theory of shadow projection. *Journal of the Warburg and Courtauld Institutes, 38,* 258–287.

Crow, F. (1977). Shadow algorithms for computer graphics. In *Siggraph '77: Proceedings of the 4th annual conference on computer graphics and interactive techniques* (pp. 242–248). New York, NY: ACM Press.

Cucchiara, R., Grana, C., Neri, G., Piccardi, M., & Prati, A. (2001). The sakbot system for moving object detection and tracking. In *Video-based surveillance systems: Computer vision and distributed processing (Part II—detection and tracking)* (pp. 145–158). New York, NY: Kluwer Academic Publishers.

Cucchiara, R., Piccardi, M., & Prati, A. (2003). Detecting moving objects, ghosts and shadows in video streams. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 25*(10), 1337–1342.

Durou, J.-D., Falcone, M., & Sagona, M. (2008). Numerical methods for shape-from-shading: A new survey with benchmarks. *Computer Vision and Image Understanding, 109*(1), 22–43.

Elder, J., Trithart, S., Pintilie, G., & MacLean, D. (2004). Rapid processing of cast and attached shadows. *Perception, 33*(11), 1319–1338.

Enns, J. T., & Rensink, R. A. (1990). Influence of scene based properties on visual search. *Science, 247,* 721–723.

Farid, H., & Bravo, M. (2010). Image forensic analyses that elude the human visual system. In *Proc. of SPIE Symposium on Electronic Imaging* (Vol. 7541), San Jose, CA.

Finlayson, G. D., Fredembach, C., & Drew, M. S. (2007). Detecting illumination in images. In *IEEE Proc. International Conference on Computer Vision (ICCV)* (pp. 1–8). Rio de Janeiro, Brazil, October 2007.

Finlayson, G. D., Hordley, S. D., Drew, M. S., & Lu, C. (2006). On the removal of shadows from images. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 28*(1), 59–68.

Fiorani, F. (2008). The colors of Leonardo's shadows. *Leonardo, 41*(3), 271–278.

Fitzpatrick, P., & Torres-Jara, E. (2004). The power of the dark side: using cast shadows for visually-guided touching. *Proc. of the 4th IEEE/RAS International Conference on Humanoid Robots* (pp. 437–449).

Galfano, G., & Pavani, F. (2005). Long-lasting capture of tactile attention by body shadows. *Experimental Brain Research, 166*(3–4), 518–527.

Hasenfratz, J.-M., Lapierre, M., Holzschuch, N., & Sillion, F. (2003, Dec). A survey of real-time soft shadows algorithms. *Computer Graphics Forum, 22*(4), 753–774.

Hering, E. (1878). *Zur lehre vom lichtsinne: Sechs mittheilungen an die kaiserl* (2d ed. ed.). Akademie der Wissenschaften in Wien: C. Gerold's Sohn.

Hoprasert, T., Harwood, D., & Davis, L. S. (1999). A robust background subtraction and shadow detection. In *Proc. Frame Rate Workshop, International Conference on Computer Vision.* Kerkyra, Corfu, Greece, September 1999.

Hsieh, J.-W., Hu, W.-F., Chang, C.-J., & Chen, J-S. (2003). Shadow elimination for effective moving object detection by Gaussian shadow modeling. *Image and Vision Computing, 21*(6), 505–516.

Huang, J. B., & Chen, C. S. (2009). Moving cast shadow detection using physics-based features. In *IEEE Proc. Computer Vision and Pattern Recognition (CVPR)* (pp. 2310–2317). Miami, FL, USA, June 2009.

Imura, T., & Tomonaga, M. (2003). Perception of depth from shading in infant chimpanzees (pan troglodytes). *Animal Cognition, 6*(4), 253–258.

Imura, T., & Tomonaga, M. (2009). Moving shadows contribute to the corridor illusion in a chimpanzee (pan troglodytes). *Journal of Comparative Psychology, 123*(3), 280–286.

Jacobson, J., & Werner, S. (2004). Why cast shadows are expendable: insensitivity of human observers and the inherent ambiguity of cast shadows in pictorial art. *Perception, 33*(11), 1369–1383.

Joshi, A., & Papanikolopoulos, N. (2008a). Learning of moving cast shadows for dynamic environments. In *Proc. of the IEEE International Conference on Robotics and Automation* (pp. 987–992).

Joshi, A., & Papanikolopoulos, N. (2008b). Learning to detect moving shadows in dynamic environments. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 30*(11), 2055–2063.

Kersten, D., Knill, D., Mamassian, P., & Bülthoff, I. (1996). Illusory motion from shadows. *Nature,* 379, p. 31.

Kersten, D., Mamassian, P., & Knill, D. (1994). *Moving cast shadows and the perception of relative depth* (Technical Report No. 6). Max-Planck-Institut für biologische Kybernetik.

Kriegman, D. J., & Belhumeur, P. N. (1998). What shadows reveal about object structure. In *ECCV '98: Proc. of the 5th European Conference on Computer Vision—Volume II* (pp. 399–414). London, UK: Springer-Verlag.

Kunii, Y., & Gotoh, T. (2003). Evaluation of Shadow Range Finder: SRF for Planetary Surface Exploration. In *Proc. of the IEEE International Conference on Robotics and Automation* (Vol. 2, pp. 2573–2578).

Lee, J.-S., Roh, S.-G., Kim, D. W., Moon, H., & Choi, H. R. (2009). In-pipe robot navigation based upon the landmark recognition system using shadow images. In *IEEE International Conference on Robotics and Automation (ICRA)* (pp. 1857–1862).

Liu, Z., Huang, K., Tan, T., & Wang, L. (2007). Cast shadow removal combining local and global features. In *IEEE Proc. Computer Vision and Pattern Recognition (CVPR)* (pp. 1–8).

Mamassian, P. (2004). Impossible shadows and the shadow correspondence problem. *Perception, 33*(11), 1279–1290.

Mamassian, P., Knill, D. C., & Kersten, D. (1998). The perception of cast shadows. *Trends in Cognitive Sciences, 2*(8), 288–295.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York, NY, USA: Henry Holt and Co., Inc.

Martel-Brisson, N., & Zaccarin, A. (2007). Learning and removing cast shadows through a multidistribution approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 29*(7), 1134–1146.

Martel-Brisson, N., & Zaccarin, A. (2008). Kernel-based learning of cast shadows from a physical model of light sources and surfaces for low-level segmentation. In *IEEE Proc. Computer Vision and Pattern Recognition (CVPR)* (pp. 1–8), Anchorage, AK.

Maxwell, B., Friedhoff, R., & Smith, C. (2008). A bi-illuminant dichromatic reflection model for understanding images. In *IEEE Proc. Computer Vision and Pattern Recognition (CVPR)* (pp. 1–8), Anchorage, AK.

Mikic, I., Cosman, P., Kogut, G., & Trivedi, M. (2000). Moving shadow and object detection in traffic scenes. In *Proc. International Conference on Pattern Recognition (ICPR)* (pp. 321–324). Barcelona, Spain.

Nadimi, S., & Bhanu, B. (2004). Physical models for moving shadow and object detection in video. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 26*(8), 1079–1087.

Ni, R., Braunstein, M., & Andersen, G. J. (2004). Perception of scene layout from optical contact, shadows and motion. *Perception, 33*(11), 1305–1318.

Ostrovsky, Y., Cavanagh, P., & Sinha, P. (2005). Perceiving illumination inconsistencies in scenes. *Perception, 34*(11), 1301–1314.

Pavani, F., & Castiello, U. (2004). Binding personal and extrapersonal space through body shadows. *Nature Neuroscience, 7*(1), 13–14.

Plato. (360 BC). *The republic*. Penguin Classics. (Translated by H.D.P. Lee, Desmond Lee, 2007).

Porikli, F., & Thornton, J. (2005). Shadow flow: A recursive method to learn moving cast shadows. In *IEEE Proc. International Conference on Computer Vision (ICCV)* (pp. 891–898).

Prati, A., Mikic, I., Trivedi, M., & Cucchiara, R. (2003). Detecting moving shadows: Algorithms and evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 25*(7), 918–923.

Renno, J. R. R., Orwell, J., Thirde, D. J., & Jones, G. A. (2004). Shadow classification and evaluation for soccer player detection. In *Proc. British Machine Vision Conference (BMVC)* (pp. 839–848). Kingston upon Thames, UK.

Rensink, R. A., & Cavanagh, P. (2004). The influence of cast shadows on visual search. *Perception, 33*(11), 1339–1358.

Rittscher, J., Kato, J., Joga, S., & Blake, A. (2000). A probabilistic background model for tracking. In *Proc. European Conference on Computer Vision (ECCV)* (pp. 336–350).

Salvador, E., Cavallaro, A., & Ebrahimi, T. (2004). Cast shadow segmentation using invariant color features. *Computer Vision and Image Understanding (CVIU), 95*(2), 238–259.

Santos, P. E., Dee, H. M., & Fenelon, V. (2009). Qualitative robot localisation using information from cast shadows. In *IEEE International Conference on Robotics and Automation (ICRA)* (pp. 220–225). Kobe, Jp.

Sattler, M., Sarlette, R., Mücken, T., & Klein, R. (2005). Exploitation of human shadow perception for fast shadow rendering. In *APGV '05: Proceedings of the 2nd symposium on Applied perception in graphics and visualization* (pp. 131–134). New York, NY, USA: ACM.

Stauder, J., Mech, R., & Ostermann, J. (1999). Detection of moving cast shadows for object segmentation. *IEEE Transactions on Multimedia, 1*(1), 65–76.

Tompkins, P., Stentz, A., & Whittaker, W. L. (2001, June). Automated surface mission planning considering terrain, shadows, resources and time. In

Proceedings of the 6th international symposium on artificial intelligence, robotics and automation in space (i-SAIRAS '01), Montreal, Canada.

Troccoli, A., & Allen, P. K. (2004). A shadow based method for model registration. In *Proc. of conference on computer vision and pattern recognition workshop (cvprw'04)* (Vol. 11, p. 169). Washington, DC.

Vazquez, E., Weijer, J. van de, & Baldrich, R. (2008). Histogram-based image segmentation in the presence of shadows and highlights. In *European conference on computer vision* (pp. 1–14). Marseille, France.

Vinci, L. da. (1888). *Notebooks of leonardo da vinci*. Project Gutenberg. (J. P. Richter translation, http://www.gutenberg.org/etext/ab 5000)

Waltz, D. (1975). Understanding line drawings of scenes with shadows. In *The psychology of computer vision* (pp. 19–91). New York: McGraw-Hill.

Wang, Y., Loe, K.-F., Tan, T., & Wu, J.-K. (2005). A dynamic Hidden Markov Random Field Model for foreground and shadow segmentation. In *Proc. IEEE Workshop on Applications of Computer Vision* (Vol. 1, pp. 474–480).

Wang, Y., Loe, K. F., & Wu, J. K. (2006). A dynamic conditional random field model for foreground and shadow segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 28,* 279–289.

Wang, Y., & Ye, G. (2008). Joint random fields for moving vehicle detection. In *Proc. British Machine Vision Conference (BMVC).* Leeds, UK.

Woo, A., Poulin, P., & Fournier, A. (1990, November). A survey of shadow algorithms. *IEEE Computer Graphics and Applications, 10*(6), 13–32.