CEIG'09, San Sebastián, Sept. 9-11 (2009)
C. Andújar and J. LLuch (Editors)

Where are the lights? Measuring the accuracy of human vision

Jorge López-Moreno, Francisco Sangorrín, Pedro Latorre and Diego Gutierrez

Universidad de Zaragoza, Spain

Abstract

In real life, light sources are frequently not present in our view field. However human vision is able to infer the illumination just by observing its effect on visible objects (serving as lightprobes) or, inverting the idea, it is able to spot an object which is incoherently lit in a composition. These lightprobes have been used by computer algorithms in the same manner to detect lights, mimicking the human visual system (HVS). It has been proved that the presence of shadows or highlights in the lightprobe affects the accuracy of HVS, although its degree of influence remains unbeknownst until now. The present work performs a psychophysical analysis which aims to provide accurate data for light detection, perception-oriented rendering, image compositing and augmented reality.

Categories and Subject Descriptors (according to ACM CCS): Computing Methodologies [I.3.7]: Computer Graphics—3D Graphics; Computing Methodologies [I.4.10]: Image Processing and Computer Vision—Image Representation

1. Introduction

The process of perception in the human visual system is a complex interaction of multiple factors, both in the subject itself and in the observed scenario. The knowledge of its limitations has helped research into developing impacting technologies in areas such as image compression (e.g.: JPEG removes frequencies which are not easily perceived by the HVS).

Technologies like augmented reality [WS02], [ZY01], image editing [YWAC06] or image forensics [JF05], [JF07] depend in great manner in the process of detecting the lighting environment and inserting new objects relit in the same fashion as their neighbors. When there is limited information (like in a single image) and due to uncontrolled factors in the input images such as lens distortion or glare, detecting light directions cannot be absolutely precise. Light detection algorithms achieve large errors in their estimations. However, these errors might go completely unnoticed by users in an image while they are easily spotted in another.

We are therefore interested in determining an error threshold below which variations in the direction vector of the lights will not be noticed by a human observer. To this end we performed a psychophysical experiment where we analyze several factors involved in the general light detection process, while measuring their exact influence for its future use in computer applications.

The paper is organized as follows. In the following section, we review prior work related to our research. In the next, we describe in detail the test, motivation and methodology used in its design. We subsequently show the results and discuss their causes, scope and applicability. Finally we conclude and propose future lines of work.

2. Previous work

In [OCS05], the authors show that even though the visual system can easily spot an anomalously lit object in an array of *identical* objets with the *same* orientation and lit *exactly* the same [ER90]; [KR92], overall performance drops when altering orientations of the equally-lit objects. This suggests that a fast parallel pattern matching mechanism in the former case is substituted by a much slower serial search in the latter, making illumination detection a difficult task for humans.

The work of [TM83] states the poor behavior of the HVS in determining the light direction by observing a lightprobe.

DOI: 10.2312/LocalChapterEvents/CEIG/CEIG09/145-151

They stated that highlights had no effect in the estimation of the illuminant's direction. However their measures were limited to cylinders (a simple geometry which varies in only one axis) and the users were asked for the direction of light (the inverse of our case). In the same line, in [MT86] the assumption that the HVS assumes objects as diffuse was proved as wrong.

Additionally [KvDP04] showed how human perception is much better at azimuth estimates (scenario used in our test, see figure 2) than at zenith estimates (when lights are coplanar with the object and the screen plane). He also proved that when shadows are present, the shadow boundaries, a first order discontinuity in shading, increased the accuracy of HVS in detecting the light field direction.

3. The test

3.1. Motivation

As we anticipated in the introduction, there are several aspects in the process of light detection: the material of the objects, the presence of visual cues (like shadows), the position of the lights, the training of the user, etc.

Thus, in order to acquire an useful measure, the present test studies the most frequent scenarios in nature focusing in the following aspects:

- The influence of the positions of the lights. The work of [OCS05] anticipates that a greater presence of shadows (produced when the light source is behind the object) increases the accuracy of HVS. We aim to measure this influence by comparing three scenarios: lights from behind, lights in the front and a mixed situation.
- The properties of the surface material of the lightprobe. In particular we analyze how specularity in the dichromatic model (composed by diffuse and specular component) affects human perception.
- The influence of genre and professional training in the accuracy of HVS.

In the end we aim to obtain a measure of accuracy in a very general scenario: multiple light sources, multiple material properties and a complete range of light positions.

Note that in the test we precluded the presence of a very strong visual cue: shadows cast in horizontal surfaces (eg.: a floor) by the objects of the scene. There are two reasons for this exclusion: First, the subject has been studied in greater depth than other aspects by previous works, its influence has been clearly stated, and which it is more important, it is a visual cue that might not be present in many scenarios (in opposition to shading, materials or self shadowing which are ever-present features).

3.2. Methodology

The psychophysical test consists of a series of images, which have been shown to 38 users. These images were generated with a 3D modelling commercial software.

Each of the images shown to the users displays a group of eight objects (Figure 1) not symmetrical with different textures and degrees of shininess. Of which two pairs have the same material, so there are a total of six different materials in the scene. One of the objects has low-frequency texture with specular reflectance (Phong type). Two of the objects do not have any texture and its reflectance is diffuse. One of the objects shows a high frequency texture and a diffuse reflectance. Two other objects, like the first two, do not have texture but its reflectance is highly specular (Phong type). The remaining two have a checkerboard texture, one with a specular reflectance (Phong type) and the other with a lambertian or diffuse reflectance.

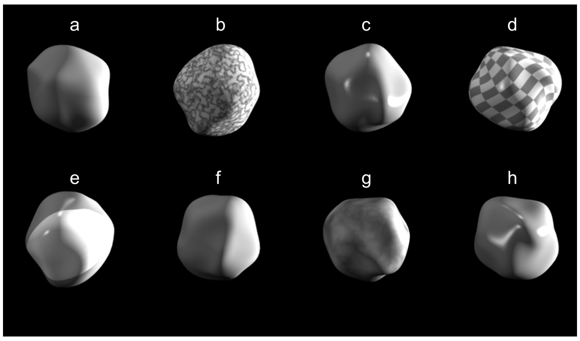

Figura 1: *One of the images shown in our test: Eight abstract objects with a main light stemming from the right.*

The shape of all the objects has the following properties and motivation:

- It is a convex shape. This is meant to be coherent with the HVS default assumption of global convexity when the humans perceive an object for first time, as proved in [LB01]. A user when asked to choose an object as lightprobe in a scene will rather choose something perceived as convex.
- It has a complex surface, with hills, valleys and creases as a real object could have. We aim to mimic a natural scene. We choose the shape following the work of [VLD07]
- It is an abstract shape, without any semantical significance for the user. We want to avoid any preferences when choosing between the objects in the scene.
- It has a different geometry than his neighbors. This is done to avoid side-by-side comparison. This kind of scenario triggers a specialized and very efficient mechanism of parallel analysis in the HVS that we should avoid in a general scenario.
- The variance in geometry between objects has to be limited, because geometry has influence in the perception of

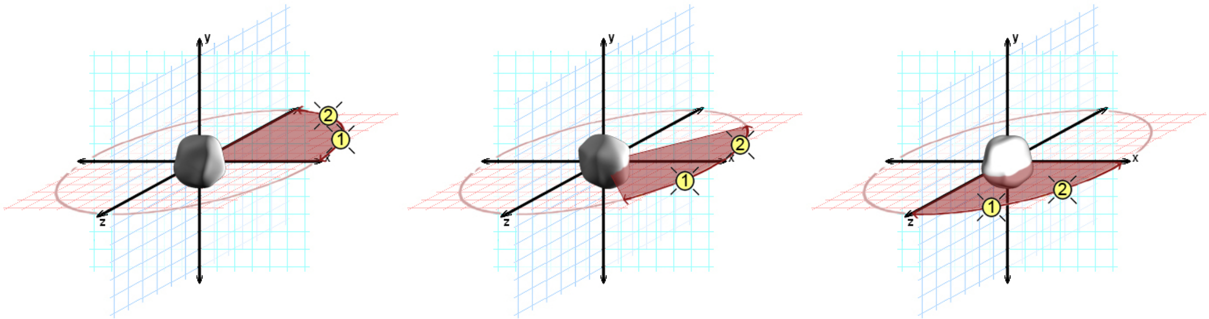

Figura 2: *3D representation of the 3 scenes (Back, Frontal-Back and Frontal) rendered in our images. The positive ϕ angle of the 2 main light sources is colored in red in the ZX plane.*

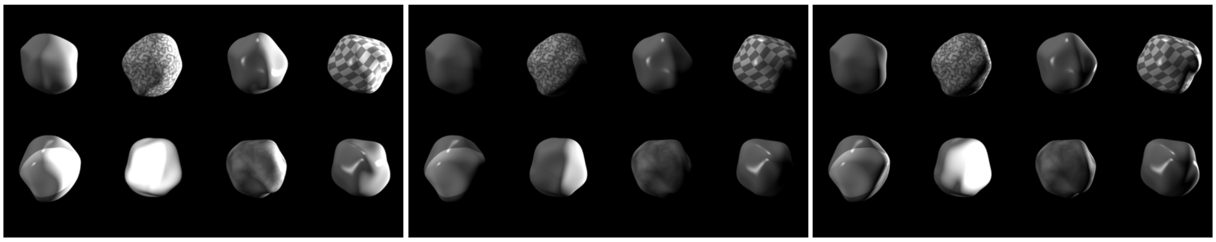

Figura 3: ***Left**: Objects illuminated by a side light ($\phi = 0°$) and a front light ($\phi = 90°$). **Middle**: Objects illuminated by a backlight ($\phi = -90°$) and a side light ($\phi = 0°$). **Right**: objects illuminated by a light with $\phi = -45°$ and a front light with $\phi = 45°$.*

reflectance [VLD07]. The objects must be visually different but also similar from a statistical point of view.

This was achieved by adding a combination of gaussian and fractal noise to geodesic spheres.

In each of the 60 images, all of the objects are illuminated with an ambient light composed by two spotlight sources, four times weaker in terms of luminance than the main light. Seven of the eight objects have as main light a directional light source, while the eighth will be subject to a light source with the same properties but different position.

The coordinate system used is the spherical coordinate system. Our two main lights (one light for 7 objects an the other ligth for the last object) will be moving along the azimuth angle ϕ of its orthogonal projection on that plane (Figure 2).

These two directional spotlights will vary its position angle along the XZ plane between different images. The absolute difference of ϕ angles between the two directional spotlights will increase from 0° to a maximum difference of 90° in steps of 10°. Thus we get ten images for each of the three different situations that we are studying in terms of light positions, which are: when both sources are illuminating the frontal hemisphere of the object, when both sources are behind the object and finally when a spotlight is on the back and another on the front (see figure 3).

In the first of the situations, the two spotlights will leave from the position of 45° between the axes +X and +Z. They will be distancing 5° each until they reach an absolute difference of 90° degrees in their ϕ angle (figure 3; Left image).

In the second situation the two lights leave from the same position, $\phi = -45°$ (between the axes +X and -Z) and they will be distancing 5° each till they align with the axis X and -Z respectively (figure 3; Middle image).

Finally, both directional lights leave from the +X axis and, as in the two previous cases they will be distancing 5° each till they reach a maximum distance of 90 degrees in their θ angle (figure 3; Right image).

To measure the influence of the material reflectance in light detection, the aforementioned process is repeated with two different materials in the image. Those two materials have no texture and a certain difference in their reflectance. One of them is rather shiny (assigned to the objects identified by letters C and H in figure 1) while the other has a diffuse behavior (objects by letters A and F).

As outlined above the total number of generated images is 60; 10 images by each of the 3 situations and the 2 materials to study. Each image has a resolution of 1024 pixels wide by 600 pixels high.

Furthermore, to prevent users from biasing the test, the images that are displayed were randomized so there is no progression in the positioning of lights between subsequent images. Finally the position of the object illuminated in a different way is not always the same, it varies in four different positions (A, F, C and H in figure 1).

We want to measure the HVS accuracy at its maximum possible efficiency, thus there will not be any time limitation to perform the task. However, to know how much time is needed by the users to achieve the maximum confidence possible in their choices, we will time each choice and users will be informed of this fact.

Concluding, each user is shown a total of 60 randomized images with unlimited time to perform the task. With these 60 images we check how capable is our human visual system of spotting illumination errors in 3 different lighting situations (both lights behind the object predominating the shadows versus the lights, one front and one back and two lights in the front, predominating the lights versus the shadows) with two different reflectances: diffuse and very shiny (Phong's model).

3.3. Administration of the test

To perform the test we have implemented a web application where the user selects the object of each scene. The test is self explanatory and designed for remote participation. However a group of 10 users was supervised under a controlled environment by one of the authors to assure that there is no bias in the unsupervised tests.

In this application we collect the following information: full name or ID, age, occupation and gender. Finally the user is asked; *Do you have any experience in art? Drawing, painting, photography, digital art* , which has to be answered affirmatively or negatively. Once the information has been collected (See figure 4) the test starts.

Figura 4: *Start page of our web application. The form collects the user information for the database.*

While users are performing the test two pieces of information are stored for each of the 60 images (figure 5). The answer selected by user (A, B, C, D, E, F, G or H) and the time it took to decide on this option. All these data are stored in a text file that later will be analyzed.

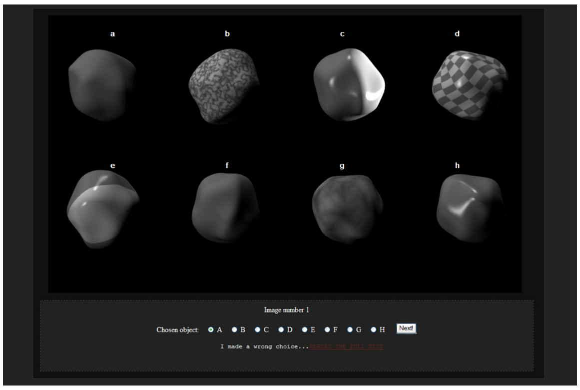

Figura 5: *Example of test image in our web application.*

4. Results and discussion

First of all we analyzed the number of correct hits depending on the degree of difference between the two lights. On one hand we analyze images in which the relit object is a diffuse (NB in figure 6) and on the other side where the object has a specular reflectance (B in figure 6). Furthermore in both graphics is included information for the three combined positions of the lights. In blue, the two are in the back (B), in garnet, at the front and back (FB) and in yellow both being frontal (F). With this information we can help to quantify how much the predominance of light or shadows affects to detecting lighting errors.

In figure 6 we can observe that up to 20 degrees the probability of detection is below chance (12.5%). In the case that both lights are in the front this probability raises up to 30 degrees. This agrees with previous studies [KvDP04] which stated that the HVS is more accurate when shadows are present.

Furthermore, the figure 7 shows the comparative analysis for objects with high specularity. We can observe that for any position of the light source, the performance of HVS is slightly lower when highlights are present. When questioned after the test, several users stated that given the number of highlights and the unknown geometry, sometimes it was difficult to decide if they were produced by the modified light or by another.

Secondly we compared the average rate of success of people with artistic experience and those who lack it. The results show a certain advantage (close to a 15%) in average for the ten subjects which claimed to have artistic skills or expertise).

We also compared the average rates by gender. There is a slightly difference between female and male subjects. We might argue that this could be due to two facts: First, our test is similar to those of mental rotation which have proved different strategies for each gender: males utilize a parietal lobule "gestalt" perceptual strategy, while females may utilize a frontal lobe "serial" reasoning strategy [HTE06].

Secondly, our scenario (timing limitation, choice of object collection) could be ill suited for the female model. For instance, works like [MNC09] show an advantage of the female analytic model in recognizing human faces, which trigger different areas of the brain than those involved when analyzing previously unknown objects as in our case. In any case, the difference observed could be significative but the influence of very complex factors in the strategies of each gender makes a difficult task of drawing solid conclusions without further studies in more scenarios.

Among the 60 images of the test we introduced 6 control images in which all objects are illuminated in the same way. These control images help us to observe the salience of the different objects. We show a bar chart with the different options that users have selected when they made a mistake (8). Each of the three bars correspond to the three positions of the lights (both lights behind the object predominating the shadows versus the lights, one front and one back and two lights in the front, predominating the lights versus the shadows).

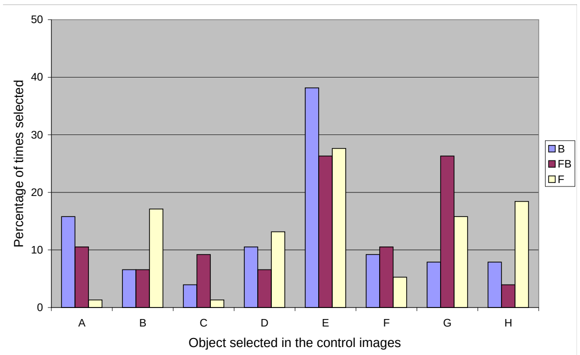

Figura 8: *Chosen object in the control images. The users have a preference for object E.*

There is an outlier with object *E* (which is not one of the diverging objects), probably due to its particular geometry and white albedo patch. In the chart of figure 9 we can observe that its salience compared with the remaining objects is reduced in direct relation with the augment of divergence. We can state that it is mistaken as the stimulus until the actual stimulus becomes noticeable in the image (around 20-30 degrees of divergence).

Regarding the time used in the decision making process, the average mark was 14.54 seconds. We have also appreciated that time of the decision process decreases as the divergence is increased.

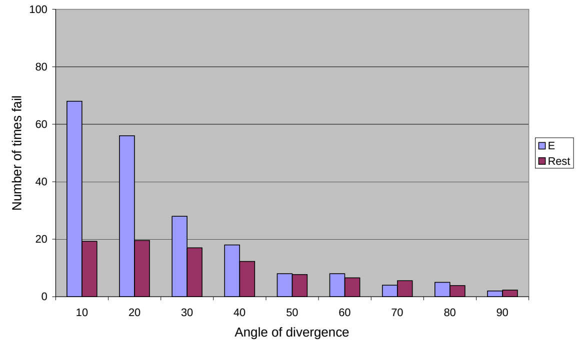

Figura 9: *The relative salience of the object E, computed as the number of times when it is chosen while missing the right choice. This is plotted in relation with the salience of the remaining objects.*

Plotting the data by material in figure 10, it is noticeable the similarity between diffuse and specular material which confirms the very limited effect of highlights in the detection of light sources. They differ only when light sources are in the back, due to the Fresnel highlights produced near the silhouette which help the user in differentiating the wrongly lit object at small divergences.

Finally in figure 11 we can observe the effects of fatigue and training in our test by studying the average time by question in the order in which the images are shown, so that when we generate a curve trend of these points we obtain the curve of fatigue of the average user. The trend line is generated with a third degree polynomial regression. The first 20 questions show a clear decrease in times due to training. After a period of stability, the last 15 questions show a new decrease, which in this case it is produced by fatigue.

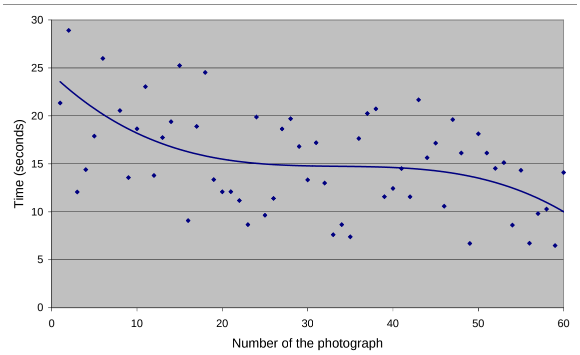

Figura 11: *Average time used by question in our test.*

5. Conclusions and future work

In the present work we have collected data from 38 users (10 of them with artistic skills) to measure the accuracy of human visual system under varying light positions an different reflectance in the objects of the scene.

Our results agree with the theories exposed in previous research on perception [OCS05], [KvDP04], furthermore we have estimated the values at which human perception is unable to spot errors in illumination in a general scenario. To our knowledge no previous research has tackled this problem.

Additionally, thanks to the data collected we can assure that a high dichromatic specularity (highlights) has little effect in the perception and we have measured that effect (in agreement with [TM83]. This suggests that the strategy of the HVS diverges form some recent computer approaches which use highlights as visual cues like [LF06].

We believe that the present work is a good start basis for those areas of computer graphics and vision which depend on analyzing the lighting environment: algorithms based in light detection and methods for image analysis and processing. By constraining their error ranges to those shown in the present paper, techniques for relighting synthetic objects to mimic their surrounding illumination (digital forgery, augmented reality,...) can work with inaccurate approximations with the certainty of being under human perception ranges.

For future research we aim to further validate our analysis with eye tracking and salience tests. Additionally, we will analyze further factors in the perception process like the spatial frequency of the textures (high frequency patterns against constant albedo values).

6. Acknowledgements

This research was partially funded by a generous gift from Adobe Systems Inc and the Spanish Ministry of Science and Technology (TIN2007-63025)

References

[ER90] ENNS J., RENSIK R.: Influence of scene-based properties on visual search. *Science 247*, 721–723 (1990).

[HTE06] HUGDAHL K., THOMSENB T., ERSLAND L.: *Applications of visual perception in computer graphics, SIGGRAPH Course 32*. No. 3. Elsevier, 2006, p. 1575Ű1583.

[JF05] JOHNSON M. K., FARID H.: Exposing digital forgeries by detecting inconsistencies in lighting. In *MM&Sec '05: Proceedings of the 7th workshop on Multimedia and security* (New York, NY, USA, 2005), ACM, pp. 1–10.

[JF07] JOHNSON M. K., FARID H.: Exposing digital forgeries in complex lighting environments. *IEEE Transactions on Information Forensics and Security 2*, 3 (2007), 450–461.

[KR92] KLEFFNER D., RAMACHANDRAN V.: On the perception of shape from shading. *Perception and Psychophysics 52* (1992), 18–36.

[KvDP04] KOENDERINK J. J., VAN DOORN A. J., PONT S. C.: Light direction from shad(ow)ed random gaussian surfaces. *Perception 33*, 12 (2004), 1405–1420.

[LB01] LANGER M. S., BÜLTHOFF H. H.: A prior for global convexity in local shape-from-shading. *Perception 30* (2001), 403–410.

[LF06] LAGGER P., FUA P.: Using specularities to recover multiple light sources in the presence of texture. In *ICPR '06: Proceedings of the 18th International Conference on Pattern Recognition* (Washington, DC, USA, 2006), IEEE Computer Society, pp. 587–590.

[MNC09] MCBAIN R., NORTON D., CHEN Y.: *Females excel at basic face perception*. No. 130. Elsevier, 2009, pp. 168–173.

[MT86] MINGOLLA E., TODD J.: Perception of solid shape from shading. *Biological Cybernetics 53* (1986), 137–151.

[OCS05] OSTROVSKY Y., CAVANAGH P., SINHA P.: Perceiving illumination inconsistencies in scenes. *Perception 34* (2005), 1301–1314.

[TM83] TODD J., MINGOLLA E.: Perception of surface curvature and direction of illumination from patterns of shading. *Journal of Experimental Psychology Human Perception and Performance 9*, 4 (1983), 583–595.

[VLD07] VANGORP P., LAURIJSSEN J., DUTRÉ P.: The influence of shape on the perception of material reflectance. *ACM Trans. Graph. 26*, 3 (2007), 77.

[WS02] WANG Y., SAMARAS D.: Estimation of multiple directional light sources for synthesis of mixed reality images. In *Proceedings of the 10th Pacific Conference on Computer Graphics and Applications* (2002), pp. 38–47.

[YWAC06] YU T., WANG H., AHUJA N., CHEN W.-C.: Sparse lumigraph relighting by illumination and reflectance estimation from multi-view images. In *Eurographics Symposium on Rendering* (2006), Eurographics Association, pp. 41–50.

[ZY01] ZHANG Y., YANG Y.-H.: Multiple illuminant direction detection with application to image synthesis. *IEEE Trans. Pattern Anal. Mach. Intell. 23*, 8 (2001), 915–920.

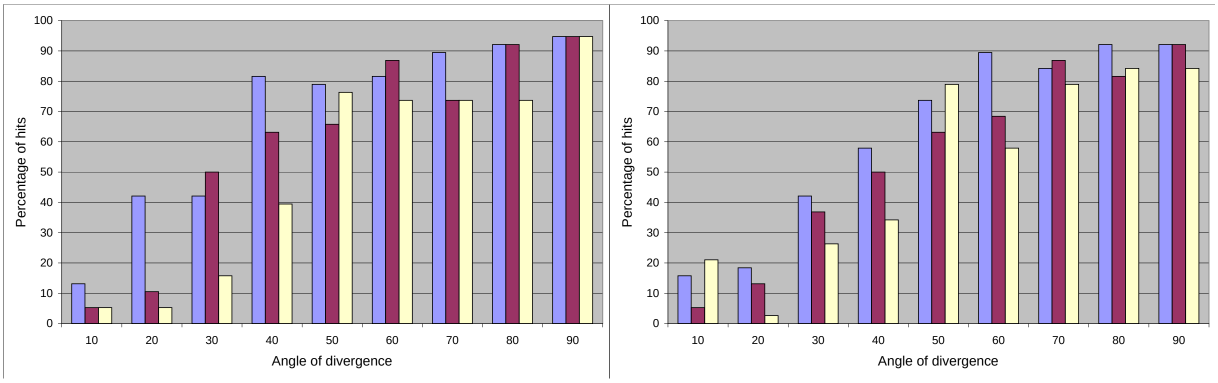

Figura 6: *Results for the three quadrants: Back (blue), Frontal-Back (red) and Frontal (yellow). The chart describes the relationship between the position of the light and the measured accuracy. Left: Hit probability with diffuse material. Right: Hit probability with a specular material.*

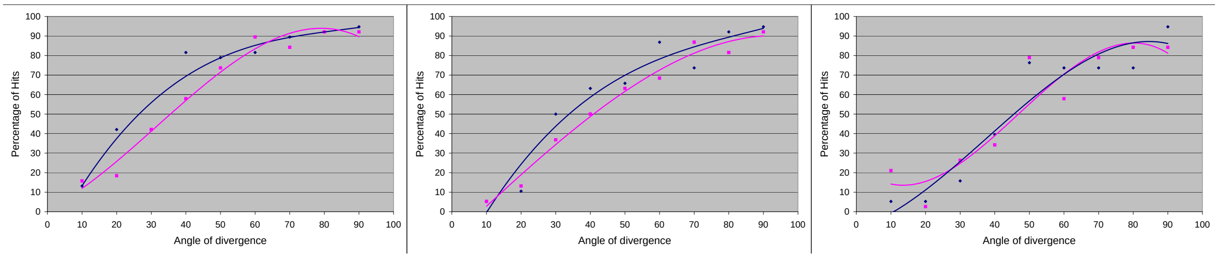

Figura 7: *Hit probability by quadrant. Left: with back position. position. Middle: with front-back position. Right: with frontal. In blue object without reflectance and in rose object with it.*

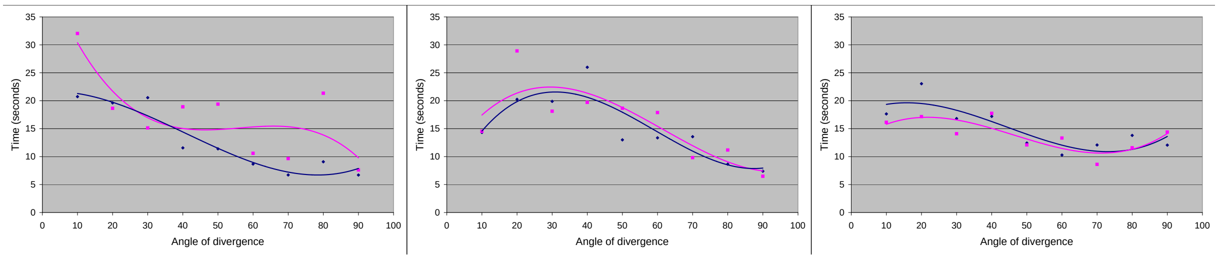

Figura 10: *Times ordered by position and material. The graphs show similar trends for both materials. Left: with back position. position. Middle: with front-back position. Right: with frontal. In blue object without reflectance and in rose object with it.*