

Aesthetics in Covariant Image Reconstruction

Todor Georgiev

tgeorgie@adobe.com

Adobe Systems, 345 Park Ave, W10-124, San Jose, CA 95110, USA

Abstract

This paper describes a method of seamless cloning based on aesthetic theory of lightness perception. Judgment of lightness and color harmony is treated as low level aesthetic judgment made by the human visual system. The equation written based on this consideration is an improvement to Poisson image editing, and produces results that are better than the current state of the art in the area of scratch/object removal. The reason our result is aesthetically pleasing is that it is fundamentally based on aesthetic theory, and it proves the usefulness of our theoretical approach.

1. Introduction

1.1. Scratch removal and image reconstruction

During the last few years there has been significant progress in the area of image reconstruction or Inpainting. The goal is to remove defects like scratches or unwanted objects by seamlessly filling in the damaged area. The state of the art today is represented by 2 algorithms, Inpainting and Poisson Editing [BSCB00, PGB03].

The method of *Inpainting* [BSCB00, BBS] is related to ideas from fluid dynamics and solves a higher order partial differential equation (PDE) to reconstruct the image in the selected area. The idea of this approach is edge continuation into the inpainting area. A simplified inpainting method that produces similar results on many pictures would be to solve Laplace's equation for the inpainting area based on Dirichlet boundary condition.

The Poisson editing approach [PGB03], based on a form of cloning, is another way of achieving high quality results. The cloning is done not in normal pixel space, but in the space of changes, also called the gradient domain [FLW02]. After cloning, the changes are integrated back to the image. That's why a better name would probably be "gradient domain cloning" or "gradient domain fusion" [ADA*04]. In this paper we call it *Poisson cloning*. Poisson cloning has been used in Adobe Photoshop since version 7.0 [Ado02, Geo04].

1.2. Reconstruction by Poisson cloning

We would like to demonstrate Poisson cloning on a typical example of an image needing repair. Figure 1 is a picture of San Marco cathedral in Venice (courtesy of Russell Williams). The image was scanned from film, with dust added on purpose.

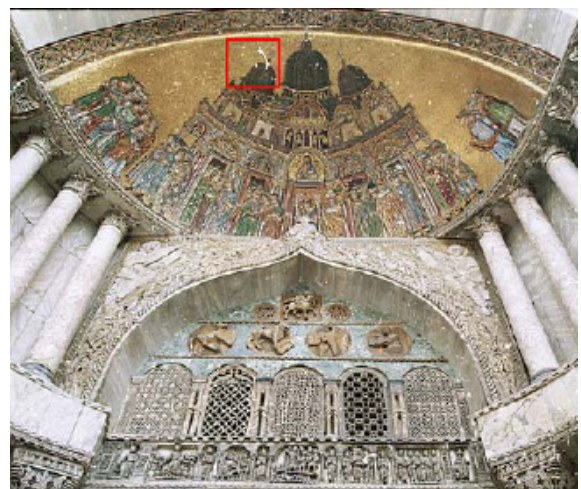


Figure 1: Basilica San Marco, Venice.

Figure 2 shows detail in the same picture. Film grain,

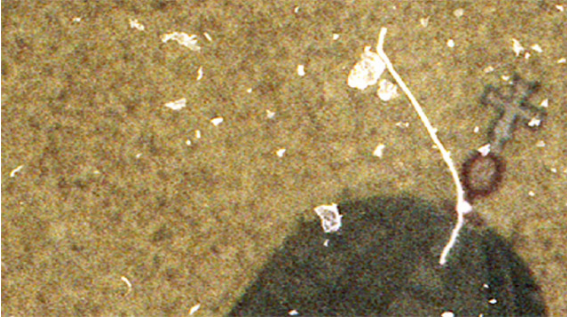


Figure 2: Detail from Figure 1.

noise and a scratch are visible. The goal is to remove the scratch in a seamless way. In Figure 3 (left) we see the result of Laplace inpainting. The method does a good job at interpolating colors in the inpainted area, but suffers aesthetically. It lacks the look and feel of real texture. It is too smooth. Adding noise is the simplest solution, often used with inpainting techniques.

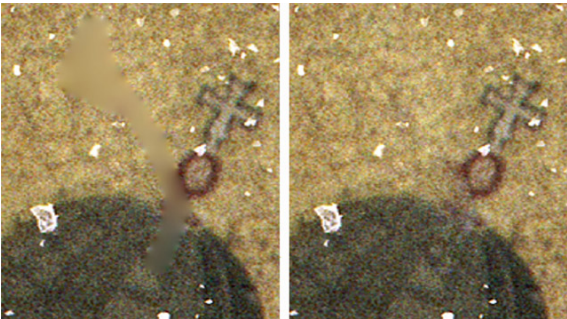


Figure 3: Scratch removed by inpainting (left) and Poisson cloning (right).

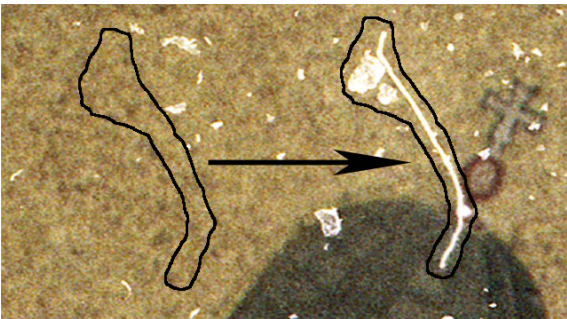


Figure 4: Areas in Figure 2 used for Poisson cloning.

Figure 3 (right) shows the scratch removed by Poisson

cloning. The source and target areas for the Poisson cloning are shown in Figure 4.

This technology was first implemented in Photoshop 7.0 [Ado02], and first described in the Poisson Image Editing paper [PGB03]. The algorithm is based on solving the Poisson equation with right hand side (source term) taken from the image in some area of texture (see Figure 4). If the grayscale image is $f(x,y)$ and the sample area image is $g(x,y)$, Poisson cloning is solving the Poisson equation

$$\Delta f(x,y) = \Delta g(x,y) \quad (1)$$

with Dirichlet boundary condition constraining the new $f(x,y)$ to match the original image at the boundary.

Everywhere in this paper

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}. \quad (2)$$

Also, $g(x,y)$ is the texture we want to transfer to the inpainted region. Texture is assumed translated to the reconstruction area.

Dirichlet boundary conditions for the Poisson equation make Poisson cloning seamlessly match the boundary of the patch.

This simple approach has been very successful, capturing a lot of attention in the media. An Internet search on Healing Brush reveals its popularity.

2. Problems with Poisson cloning

Our current paper describes an improvement to Poisson cloning based on our theory of aesthetics. Poisson cloning between areas of different lighting conditions can be a problem without this improvement.

To provide a clean example of the problem, let's try to remove the scratch from the shadow area in Figure 5 using only source material from the illuminated area.

Figure 6 shows the result of Laplace inpainting. Again, it is too smooth.

In Figure 8, we see the result of Poisson cloning from illuminated area into the shadow area. It correctly matches pixel values at the boundary of the patch, but the cloned pebbles are still easy to spot. Aesthetically we are not satisfied with the result. There is too much variation, too high contrast, in the "healed" area of the image. This problem is inherent in the nature of the Poisson equation (1), which



Figure 5: Original image of pebbles and a scratch.



Figure 6: Scratch removed by simple Laplace inpainting.

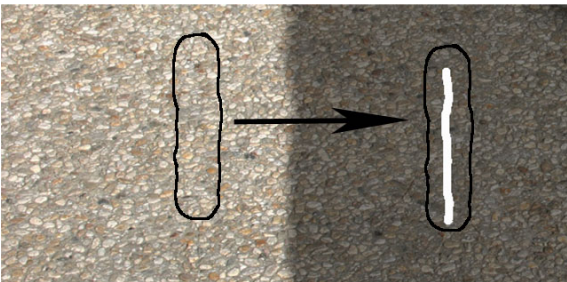


Figure 7: Areas used for Poisson cloning in Figure 8 and covariant reconstruction, Figure 11.

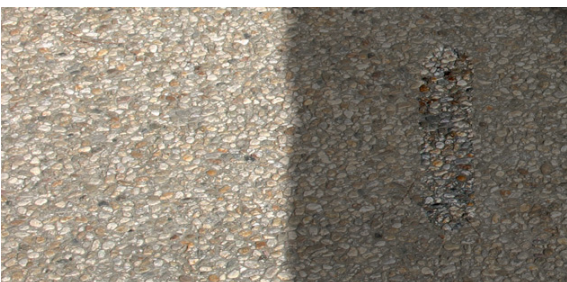


Figure 8: Scratch removed by Poisson cloning from the illuminated area.

transfers *variations* of g directly, without modifying their

amplitude even if new brightness values are modified to match the surroundings.

3. Aesthetics in low level vision

In order to solve the above problem, we want to step back and consider it from the point of view of what we call theoretical aesthetics. What are the conditions that make this possible in the first place? Where are the roots of the problem?

Human vision constantly adapts to changing lighting conditions both on global and local level. When a person looks at a scene under bright lighting, he sees an image. If the lighting is changed, human vision compensates internally for this change of lighting, and we call the process *global adaptation* to relighting.

Human visual perception also includes *local adaptation*, in which a person seeing an object partly in light and partly in shadow correctly interprets the image's color as continuous and discounts the apparent change at the shadow line. Thus the *lightness*, or perceived brightness, at any pixel in an image is determined by the context of that pixel.

Our eyes perceive lightness differently from the way a camera would capture the same pixel values. Lightness (perceived brightness) depends in a complex way on adaptation to surrounding, and previously observed pixels. Adaptation changes our fundamental judgement of lightness and color, and this can be demonstrated with a number of well known illusions [Gaz00, Sec00].

We add something to what we observe. We create a new world just by perceiving it, by adding our subjective thing to the object. A philosophical approach to this situation is related to the *form of sensibility* or *intuition* in Kant's Transcendental Aesthetic [Kan98].

In a related way, there is something fundamentally aesthetic about being able to perceive the world, as well as to perceive illusions. If we were to see things like a camera, without intuition, the world would be just a record of data, and there would be no illusions.

Color harmony, the influence of one color on another, and "simultaneous contrast" illusions are just some of the examples of what we call *aesthetics of low level vision*. Pixel comparison is fundamentally aesthetic, and not objective (physical) measurement.

Adaptation works even on the semantic level. Semantic information about the image can change perceived lightness. In Figure 9, two identical grey parallelograms p and r are perceived as different in lightness because the viewer knows that one is an illuminated top p of a dark cube q , and the other is a face r of a light cube.

The transformation of luminance into lightness can occur

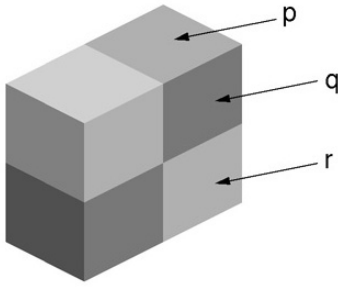


Figure 9: *Semantic adaption to lighting change.*

differently at each pixel. We can force ourselves to believe we see things one way or the other. For example if you think of Figure 9 as representing a concave shape, the illusion of difference in lightness of p and r disappears.

Our point is to put forward the idea that there is no objective or “true” perceived lightness. Judging lightness is an aesthetic process that depends on internal factors, or aesthetic criteria related to adaptation.

4. The Covariant derivative

In what follows we borrow from the Retinex [Lan77, Hor74] and the von Kries [vK02] theories of the adaptation of human vision. For the mathematical aspect of our approach, including introduction to covariant derivatives and treatment of color, we should point the reader to [Geo05b].

The image is just a record of pixel values. We do not see pixels by measuring those values directly. Instead, we perceive them through our internal adaptation. The result is not a replica of the pixel record, but an illusion, or a subjective interpretation of it.

Adaptation is what gives meaning or way of comparison to pixels. In the spirit of Kant, this is the “intuition”, the aesthetic part that creates a picture out of a meaningless pixel record.

If we express the same in mathematical terms, the above model of the image is a section in a fibre bundle [Sau89]. In simple words this is like a function, for which the scale is undefined at each point. Each point or “fibre” can have its own scale. Not having a common scale is what creates the situation of not being able to compare pixels.

Adaptation is what makes comparison or common scale possible. In mathematics this is called *connection* because it “connects” one pixel (or fibre) to another. The term used in Physics is Covariant derivative [EG96, Wey23]. We would like to call it perceptual derivative, or adapted derivative because it describes adaptation.

To make this all more intuitive, let’s look at one example. The simultaneous contrast illusion, Figure 10 shows that



Figure 10: *The central rectangle has constant pixel values.*

humans do not perceive luminance directly. (See [Gaz00, Sec00] for a general survey on lightness perception and examples of illusions.) The figure contains a constant gray band surrounded by a variable background. Due to our visual system’s adaptation to the surrounding area, the band appears to vary in *lightness* (perceived brightness) in opposition to its surroundings. Real gradient is zero, but perceived or covariant gradient is not zero. Our perception of changing lightness in the band is due to the covariant gradient being nonzero.

If the equations we use do not reflect this situation, they can not produce results that are acceptable to our visual system. This explains the negative result in Figure 8. The covariant or perceptual derivative is a useful tool for making the equations change “co-variantly” with the adaptation of the visual system (see [Geo05a]). The current paper will show how it fixes the type of problem in Figure 8.

In this paper we provide the mathematical recipe that describes effects of adaptation illustrated in Figure 10. In the usual equations we simply replace each derivative with a covariant derivative. These covariant derivatives are specified so that the covariant gradient is equal to the perceived gradient. In the example of Figure 10, constant pixel values in the band have nonzero covariant derivative and describe perceived gradient.

5. Main equations

Following the example of Electrodynamics and Quantum Mechanics, we will replace conventional derivatives with *covariant derivatives*. In our approach they describe adaptation of the visual system in the following way. Perceptually correct gradient is written based on the following recipe: Each derivative is replaced with a “derivative + function” expression:

$$\frac{\partial}{\partial x} \rightarrow \frac{\partial}{\partial x} + A_x(x, y) \quad (3)$$

$$\frac{\partial}{\partial y} \rightarrow \frac{\partial}{\partial y} + A_y(x, y) \quad (4)$$

Here A_x and A_y are the x and y components of the vector function $\mathbf{A}(x, y)$, which is used to describe the adaptation of the visual system. It represents the additional freedom which redefines our perception of gradients based on adaptation and will be specified later, in equations (8), (9) and (10). The gradient visible in Figure 10 is due to covariant derivative adapted to the surroundings.

It is well known that the Laplace equation $\Delta f = 0$ with Dirichlet boundary conditions is the simplest way to reconstruct (or inpaint) a defective area in an image. Let's write the derivatives explicitly:

$$\frac{\partial}{\partial x} \frac{\partial}{\partial x} f + \frac{\partial}{\partial y} \frac{\partial}{\partial y} f = 0, \quad (5)$$

After performing the above substitution (3), (4), the Laplace equation (5) is converted into the *covariant Laplace equation*:

$$\left(\frac{\partial}{\partial x} + A_x\right)\left(\frac{\partial}{\partial x} + A_x\right)f + \left(\frac{\partial}{\partial y} + A_y\right)\left(\frac{\partial}{\partial y} + A_y\right)f = 0, \quad (6)$$

which after differentiation can be written as

$$\Delta f + f \operatorname{div} \mathbf{A} + 2\mathbf{A} \cdot \operatorname{grad} f + \mathbf{A} \cdot \mathbf{A} f = 0. \quad (7)$$

Here the vector function $\mathbf{A}(x, y) = (A_x(x, y), A_y(x, y))$ describes adaptation of the visual system, or (which is the same) the aesthetics of perception. It is also related to the "guidance field" in Poisson Image Editing [PGB03], and it is playing the same role as the vector potential in Electrodynamics.

Here is how we define $\mathbf{A}(x, y)$ in the case of our improvement of Poisson cloning. We assume the visual system is completely adapted to the area of texture, i.e. adapted to $g(x, y)$. In other words, adaptation is such that $g(x, y)$ is covariantly constant, the covariant derivatives of g are zero.

$$\left(\frac{\partial}{\partial x} + A_x(x, y)\right)g(x, y) = 0 \quad (8)$$

$$\left(\frac{\partial}{\partial y} + A_y(x, y)\right)g(x, y) = 0 \quad (9)$$

Solving for $\mathbf{A}(x, y)$ produces the specific form of the vector function that we are going to use:

$$\mathbf{A}(x, y) = -\frac{\operatorname{grad} g}{g} \quad (10)$$

Substituting in equation (7), we obtain the final form of the covariant Laplace equation:

$$\frac{\Delta f}{f} - 2\frac{\operatorname{grad} f}{f} \cdot \frac{\operatorname{grad} g}{g} - \frac{\Delta g}{g} + 2\frac{(\operatorname{grad} g) \cdot (\operatorname{grad} g)}{g^2} = 0. \quad (11)$$

We see that the covariant Laplace equation is more complicated, and actually very different, from the Laplace equation. In a way, this is a Poisson equation with a modified Δg term on the "right hand side". However, the structure of the equation prescribed by the covariant derivatives formalism is very specific.

This section attempted to show that the expression $2\frac{\operatorname{grad} f}{f} \cdot \frac{\operatorname{grad} g}{g} + \frac{\Delta g}{g} - 2\frac{(\operatorname{grad} g) \cdot (\operatorname{grad} g)}{g^2}$ is the correct one to choose as a source term in the modified Poisson equation for seamless cloning based on the theory of adaptation.

6. Results

We solve (11) by iteration with the appropriate kernel. For details see [Geo05b, Geo05a, PTVF92] Pyramidal application of the algorithm greatly speeds up calculations. A multigrid approach with better performance is described in [PTVF92]. In practical terms the tool works sufficiently fast for using it in interactive mode. For example, on a laptop running Windows XP with a 2 GHz Pentium 4 processor, applying a brush of radius 100 pixels takes about 0.25 seconds to converge. On a wide range of typical machines, when retouching an image of several megapixels with a reasonable (for the image) size brush, the iterations converge within a fraction of a second and the user does not notice slowness or unresponsiveness.



Figure 11: Scratch removed by covariant cloning from the same illuminated area as in Figure 8.

Applied to the example of Figure 8, our approach fixes the problem by automatically modifying the contrast relative

to the surrounding shadows and shadings, just as the human visual system discounts for shadows during adaptation. See Figure 11. This behavior is intrinsic to our model based on covariant (or adapted, perceptual) derivatives, which follow directly from the theory. The proposed approach often produces good results as in Figure 11. We should note however that the results are not always perfect. Sometimes we see results that are somewhere between Figure 8 and Figure 11 in quality. But in our extensive experimentation we have not been able to find a single image where the covariant method produces result worse than Poisson Editing. In cases where Poisson Editing works well, like Figure 3 (right), the covariant approach produces almost the same, but slightly better results. Often the difference is hard to distinguish.

In the end we should note that what's really important is not so much the result of scratch removal, but the perceptual and aesthetical method of covariant derivative. The results show that the method is of real value, and not just theory.

7. Conclusion and future work

The covariant (or adapted) derivative provides a way to perform perceptual image processing. It treats images according to how they are perceived as opposed to the traditional way of treating images according to how they are recorded by the camera. In this way our approach is more human or aesthetically correct, equivalent to image processing of the internal mental image inside the brain. The algorithm is based on aesthetic theory of perception, where measurement is replaced by aesthetic judgement. Gradient domain image processing is extended into covariant (perceived) gradient domain image processing.

References

- [ADA*04] AGARWALA A., DONTCHEVA M., AGRAWALA M., DRUCKER S., COLBURN A., CURLESS B., SALESIN D., COHEN M.: Interactive digital photomontage. *ACM Transactions on Graphics* 23, 3 (Aug. 2004), 294–302. 1
- [Ado02] ADOBE: *Photoshop 7.0 User Guide*. Adobe Systems, San Jose, 2002. 1, 2
- [BBS] BERTALMIO M., BERTOZZI A. L., SAPIRO G.: Navier-stokes, fluid dynamics, and image and video inpainting. *Proceedings of the International Conference on Computer Vision and Pattern Recognition, IEEE, Dec. 2001, Kauai, HI, volume 1, pp. 1-355-1362*. 1
- [BSCB00] BERTALMIO M., SAPIRO G., CASELLES C., BALLESTER C.: Image inpainting. In *Siggraph 2000, Computer Graphics Proceedings* (2000), Akeley K., (Ed.), Annual Conference Series, ACM Press / ACM SIGGRAPH / Addison Wesley Longman, pp. 417–424. 1
- [BVSO03] BERTALMIO M., VESE L., SAPIRO G., OSHER S.: Simultaneous structure and texture image inpainting. In *IEEE Transactions on Image Processing*, 12(8) (2003), pp. 882–889.
- [CAK03] CRIMINISI A. P. P., K. T.: Object removal by exemplar-based inpainting. In *Proceedings of CVPR 2003, vol. 2* (Madison, Wisconsin, June 2003), pp. 721–728.
- [CBDM82] CHOQUET-BRUHAT Y., DEWITT-MORETTE C.: *Analysis, Manifolds and Physics*. Amsterdam, North-Holland, 1982.
- [EG96] EINSTEIN A., GROSSMANN M.: *Outline of a Generalized Theory of Relativity and of a Theory of Gravitation*. Princeton University Press, 1996. ISBN 0-691-02610-6. 4
- [EL99] EFROS A. A., LEUNG T. K.: Texture synthesis by non-parametric sampling. In *IEEE International Conference on Computer Vision* (Corfu, Greece, September 1999), pp. 1033–1038.
- [FBM*90] FIORENTINI A., BAUMGARTNER G., MAGNUSSEN S., SCHILLER P. H., THOMAS J. P.: The perception of brightness and darkness relations to neuronal receptive fields. In *Visual perception the neurophysiological foundations*, Spillmann L., Werner J. S., (Eds.). Academic Press, 1990, ch. 7, pp. 129–161.
- [FHD02] FINLAYSON G., HORDLEY S., DREW M.: Removing shadows from images, 2002.
- [FLW02] FATTAL R., LISCHINSKI D., WERMAN M.: Gradient domain high dynamic range compression. *ACM Transactions on Graphics* 21, 3 (July 2002), 249–256. 1
- [Gaz00] GAZZANIGA M.: *The New Cognitive Neurosciences (2nd ed.)*. MIT Press, Cambridge, MA, 2000. 3, 4
- [Geo04] GEORGIEV T.: Photoshop healing brush: a tool for seamless cloning. In *Workshop on Applications of Computer Vision in conjunction with ECCV 2004* (Prague, May 2004), pp. 1–8. 1
- [Geo05a] GEORGIEV T.: Image reconstruction invariant to relighting. In *Eurographics 2005* (Dublin, August 2005), pp. 61 – 64. 4, 5
- [Geo05b] GEORGIEV T.: Vision, healing brush, and fiber bundles. In *Proceedings of SPIE Vol. 5666* (San Jose, January 2005). 4, 5
- [Hor74] HORN B. K. P.: Determining lightness from an image. *Computer Graphics and Image Processing* 3, 1 (Dec. 1974), 277–299. 4
- [HT96] HIRANI A. N., TOTSUKA T.: Combining frequency and spatial domain information for fast interactive image noise removal. In *SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1996), ACM Press, pp. 269–276.

- [Kan98] KANT I.: *Critique of Pure Reason*. Cambridge Univ. Press, Cambridge, 1998. 3
- [Lan77] LAND E. H.: The retinex theory of color vision. *Scientific American* 237, 6 (Dec. 1977), 108–128. 4
- [LLW04] LEVIN A., LISCHINSKI D., WEISS Y.: Colorization using optimization. *ACM Transactions on Graphics* 23, 3 (Aug. 2004), 689–694.
- [LZPW03] LEVIN A., ZOMET A., PELEG S., WEISS Y.: Seamless image stitching in the gradient domain, 2003.
- [PGB03] PÉREZ P., GANGNET M., BLAKE A.: Poisson image editing. *ACM Transactions on Graphics* 22, 3 (July 2003), 313–318. 1, 2, 5
- [PTVF92] PRESS W. H., TEUKOLSKY S. A., VETTERLING V., FLANNERY B. P.: *Numerical Recipes in C: The Art of Scientific Computing (2nd ed.)*. Cambridge University Press, Cambridge, 1992. ISBN 0-521-43108-5. 5
- [Sau89] SAUNDERS D. J.: *The geometry of jet bundles*. Cambridge Univ. Press, Cambridge, 1989. 4
- [Sec00] SECKEL A.: *The Art of Optical Illusions*. Carlton Books, 2000. 3, 4
- [vK02] VON KRIES J.: *Chromatic Adaptation, Translation: D. MacAdam, Sources of Color Science, MIT press 1970*. MIT press, 1902. 4
- [Wey23] WEYL H.: *Gravitation and Electricity*. Methuen and Co. LTD., London, 1923. 4
- [WL00] WEI L.-Y., LEVOY M.: Fast texture synthesis using tree-structured vector quantization. In *Siggraph 2000, Computer Graphics Proceedings (2000)*, Akeley K., (Ed.), Annual Conference Series, ACM Press / ACM SIGGRAPH / Addison Wesley Longman, pp. 479–488.