Midway image equalization

Julie Delon
CMLA, ENS Cachan, 61 av. du président Wilson, 94235 Cachan cedex,France
delon@cmla.ens-cachan.fr

Abstract

Midway image equalization means any method giving to a pair of images the same histogram, while maintaining as much as possible their previous gray level dynamics. In this paper, we present an axiomatic analysis of image equalization which leads us to derive two possible methods. Both methods are then compared in theory and in practice for two reliability criteria, namely their effect on quantization noise and on the support of the Fourier spectrum. A mathematical analysis of the properties of the equalization methods is performed. Their algorithms are described and they are tested on such typical pairs as satellite image stereo pairs and different photographs of a same painting.

1 Introduction

The comparison of two images, in order to extract a mutual information, is one of the main themes in computer vision. The pair of images can be obtained in many ways : they can be a stereo pair, two images of the same object (a painting for example), multi-channel images of the same region, images of a movie, etc. This comparison is perceptually greatly improved if both images have the same grey level dynamics. In addition, a lot of image comparison algorithms, based on gray level, take as basic assumption that intensities of corresponding points in both images are equal. It has, however, been often emphasized that the constant brightness hypothesis is nothing but a visual illusion [4]. As is well known by experts in stereo vision, this assumption is generally false for stereo pairs and deviations from this assumption cannot even be modelled by affine transforms [1]. Consequently, if we want to compare visually and numerically two images, it is useful to give them first the same dynamic range and luminance. We will limit the study to gray level images (the case of color images can be deduced via their intensity components).

We are led to the following problem.

- From two images I_1 and I_2, construct by contrast changes two images $\tilde{I}_1$ and $\tilde{I}_2$, which have the same histogram.
- This common histogram h should stand midway between the previous histograms of I_1 and I_2, and as close as possible to each of them. This treatment must avoid to favor one histogram rather than the other.

With this intention, two additional aspects are taken into consideration :

- Digital images being quantized, the quantization noise must remain as uniform as possible.
- I_1 and I_2 must remain "well sampled" if they were, namely, the support of their Fourier spectrum should not expand too much by the applied contrast change.

Obviously, the involved contrast changes have to be increasing and continuous, for reasons of reversibility and coherence. We will call any method addressing the problem with the above mentioned requirements a *midway image equalization*.

As we shall see, the requirements yield naturally two methods. The first one is to choose the midway histogram as an "inverse mean" of the previous ones. We will see how this method catches up with the work of Cox et al. [1], who addressed the same problem and proposed a discrete algorithm. The second one we shall derive is inspired by *histogram specification*. It consists, if H_1 and H_2 are the cumulative histograms of I_1 and I_2, to find a contrast change ψ such that $\psi \circ \psi = {H_1}^{-1} \circ H_2$ and to replace I_2 by $\psi(I_2)$ and I_1 by $\psi^{-1}(I_1)$. This second method is more difficult to approach, because there exist several solutions to the "*square root*" problem, and their graphs can be, as we shall see, oscillatory. We will introduce and describe the two announced methods in the next section. The third part of the paper deals with the theoretical construction and mathematical properties of the *square iterative root*. We will see how to choose this square root optimally towards the quantization noise. The final section will be devoted to the numerical experiences, to the comparison of both methods with specification on several examples, and their effects on the spectrum of the images and quantization noise.

2 How to equalize two histograms ?

2.1 Definitions and properties

Let $I : \Omega \to [0,1]$ [1] ($\Omega \subseteq \mathbb{R}^2$) be an image and h the distribution of its values (also called histogram of I). The *cumulative histogram* of I is the function $H : [0,1] \to [0,1]$ defined by :

$$H(x) = \int_0^x h(t)dt$$

Proposition 1 *Let $\varphi : [0,1] \to [0,1]$ be a continuous and strictly increasing function (such a function will be called an increasing homeomorphism on $[0,1]$). Then $\varphi(I)$ is a new image, whose cumulative histogram is $H \circ \varphi^{-1}$.*

We point out that such a transformation preserves the order of the gray levels in I and doesn't create or cancel any image information.
If we assume that H is continuous and strictly increasing, it follows that $H(I)$ is an image whose distribution function is uniform on $[0,1]$. This transformation is called *histogram equalization*. General information on histogram modification can be found, for example, in [2].

2.2 A description of classical methods.

Consider two images I_1 and I_2, whose cumulative histograms are H_1 and H_2. In order to make them have the same histogram, we have to choose a cumulative histogram H, and to make them correspond to H by a contrast change. In the light of the above proposition, we see that the only way to perform such a transformation is to replace each image I_i by $\varphi_i(I_i)$ where $\varphi_i = H^{-1} \circ H_i$. All histogram equalization methods are based on this principle.

1. **"Classical equalization" method** H is the identity function on $[0,1]$. Each image I_i becomes $H_i(I_i)$, i.e. each histogram is equalized. This method is easy to perform. However, if both histograms are similar and far from being constant, it is not judicious to flatten them to make them correspond. Indeed, the contrast change, used to transform the histogram of an image, affects the aspect of this image, all the more when its first histogram is far from the flat histogram to reach. As is well known and easy to check, the effect of this method is perceptually bad because of the enhancement of quantization noise (contouring artifacts).

2. **Specification** One of the images is chosen as a reference, and H is taken as the cumulative histogram of this reference. For example, if I_1 is the reference, I_1 remains unchanged and

[1] This definition is not restrictive, the range of an image can be put on $[0,1]$ by an affine transformation.

I_2 becomes ${H_1}^{-1} \circ H_2(I_2)$, which has the same histogram as I_1 (we say that h_2 is *specified* on h_1). This method is called *histogram specification* [2]. It can be judicious if both images have the same kind of dynamic range. Only one histogram is modified. For the same reason as in equalization, this method can also product contouring artifacts. A little alteration of each image should be less affecting than a big modification of one of them.

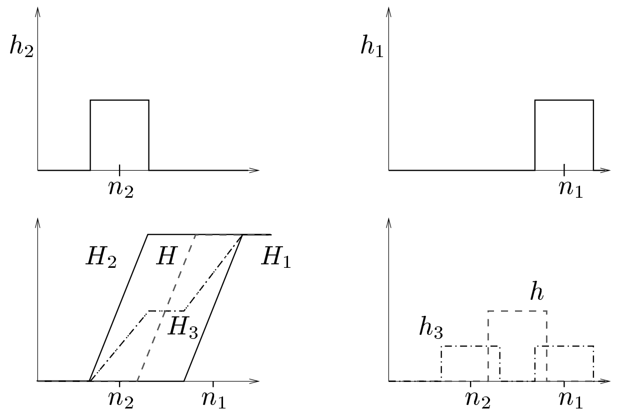

Figure 1: h_1 and h_2 are the histograms of two images I_1 and I_2. h_3 is the direct average of h_1 and h_2, but h seems to be the most natural compromise between them.

3. Mean If we want H as a compromise between H_1 and H_2, the simplest function which we can think of, is the average of those two cumulative histograms. However, the following example proves that this idea is not judicious : let I_1 and I_2 be two images, whose histograms h_1 and h_2 are represented at the top of the Fig.1. The images have no common grey level. Now, their cumulative histograms H_1 and H_2 are similar, being two step functions, respectively centered at n_1 and n_2. It follows than the most natural intermediary (in the sense stated in the introduction) for them should be a step function, placed exactly in the middle of H_1 and H_2 (H is that histogram). Now, let H_3 be the average of the cumulative histograms H_1 and H_2. We see that H_3 is made up of two successive steps instead of one. Actually, h_3 is not a good compromise between h_1 and h_2. Indeed, the composition of I_1 by ${H_3}^{-1} \circ H_1$ spreads out the grey levels of the image on the two regions centered at n_1 and n_2, in order to make its histogram correspond as much as possible to h_3. This "stretching" has at least one drawback : the grey levels of I_1, which were uniformly distributed around n_1, are separated in two distant regions, and this separation creates some structure in I_1 which didn't appear before. If we use H instead of H_3, this phenomenon doesn't appear : the grey levels of I_1 and I_2 are just translated in order to be distributed around $\frac{n_1+n_2}{2}$.

2.3 Axioms of midway equalization

We keep the notations of the previous section. We search for two contrast changes φ_1 and φ_2 such that $\varphi_1(I_1)$ and $\varphi_2(I_2)$ have the same histogram. Thus,

$$ {\varphi_1}^{-1} \circ \varphi_2 = {H_1}^{-1} \circ H_2. $$

In order to select the right method, two axiomatic methods are possible : a first one would enhance invariance and conservation principles for all images. Now, another more intuitive way to specify axiomatically a method is to state what the expected result should be for a short list of key examples. Here, we choose this second axiomatic specification method by considering two very simple particular cases : the case when I_1 and I_2 differ by a constant and the case where I_1 and I_2 indeed derive from the same image up to a contrast change, namely $I_1 = f(I)$ and $I_2 = g(I)$.

Axiom 1 : If $I_2 = I_1 + \lambda$, then $\varphi_1(I_1) = \varphi_2(I_2) = I_1 + \frac{\lambda}{2}$.

In the example presented in the previous section (Fig.1), this axiom yields exactly the only sound midway histogram. Assume now that I_1 and I_2 represent the same image, i.e. that there exists an image I and two increasing functions f and $g : [0,1] \to [0,1]$ such that $I_1 = f(I)$ and $I_2 = g(I)$. We shall consider two alternative axioms, both natural in this context.

Axiom 2 : in the case $f = g^{-1}$, the intermediate image between I_1 and I_2 is I

or,

Axiom 3 : the intermediate image between I_1 and I_2 is a direct average $F(f(I), g(I))$ where $F : [0,1]^2 \to [0,1]$ is a real function, independent of f and g.

This axiom means that there exists a function F such that for every pair of images (I_1,I_2), the inverse of their intermediate cumulative histogram will be $F(H_1^{-1}, H_2^{-1})$.

We shall now examine the pairs of axioms 1 & 2, and 1 & 3 in turn. Each pair leads to a different midway equalization method.

Analysis of Axioms 1 & 2 : the Square Root method. Notice that axiom 2 implies the axiom 1. Indeed, if $I_2 = I_1 + \lambda$, we can write $I_2 = (I_1 + \frac{\lambda}{2}) + \frac{\lambda}{2}$ and $I_1 = (I_1 + \frac{\lambda}{2}) - \frac{\lambda}{2}$. Then the intermediate image is $I_1 + \frac{\lambda}{2}$. The choice of the axioms 1 and 2 leads to construct as a new histogram $H_{new} = H \circ f$ (if H is the cumulative histogram of I). The function f is unknown, but verifies the equality

$$H_1{}^{-1} \circ H_2 = (H \circ f^{-1})^{-1} \circ (H \circ g^{-1}) = f \circ f$$

Thus, the general method following this case consists precisely in building an increasing function ψ such that $\psi \circ \psi = H_1{}^{-1} \circ H_2$ and replacing I_2 by $\psi(I_2)$ and I_1 by $\psi^{-1}(I_1)$ (i.e. we assume $\varphi_1{}^{-1} = \varphi_2$). In the next section, we will see how to construct such a function. We can argue that if such a *square iterative root* exists, it will be closer to the identity than the function $H_1{}^{-1} \circ H_2$. Consequently, the changes performed on both images by using ψ or ψ^{-1} will be smaller than the one which is performed on only one image by specification.

Analysis of Axioms 1 & 3 : the Inverse Average Equalization. If the axioms 1 and 3 are satisfied, the axiom 1 implies that :

$$\forall x,\ \forall \lambda,\ F(x + \lambda, x) = x + \frac{\lambda}{2},$$

which yields $F(x,y) = \frac{x+y}{2}$, $\forall (x,y) \in [0,1]^2$. Thus, we have :

Proposition 2 *the only intermediate image between $f(I)$ and $g(I)$ satisfying the axioms 1 and 3 is $\frac{f(I)+g(I)}{2}$.*

The intermediate cumulative histogram constructed in this way is

$$\begin{aligned} H_{new} &= H \circ (\frac{f+g}{2})^{-1} = (\frac{f \circ H^{-1} + g \circ H^{-1}}{2})^{-1} \\ &= (\frac{H_1{}^{-1} + H_2{}^{-1}}{2})^{-1}. \end{aligned}$$

The general method deduced from the preceding proposition will be called *Inverse Average Equalization* (i.a.e.). The contrast changes involved are :

$$\varphi_1 = \frac{id + H_1{}^{-1} \circ H_2}{2} \quad \text{and} \quad \varphi_2 = \frac{id + H_2{}^{-1} \circ H_1}{2}$$

The Cox et al. algorithm. We shall now prove that the Inverse Average Equalization is nothing but a continuous version of the discrete algorithm proposed by Cox et al. in [1]. The dynamic histogram warping (DHW) presented in [1] uses dynamic programming for matching directly histogram values. Only one-to-one, one-to-many or many-to-one mappings are allowed.

Suppose the range $[n_1, \tilde{n_1}]$ in I_1 has been mapped with the intensity n_2 in I_2. Cox et al. propose to simultaneously warp both histograms by replacing the previous grey levels in both images by [2] $n_0 = \frac{n_2 + \frac{n_1+\tilde{n_1}}{2}}{2}$. Now, we have $H_1(\frac{n_1+\tilde{n_1}}{2}) \simeq H_2(n_2)$, then :

$$n_0 \simeq \frac{n_2 + {H_1}^{-1} \circ H_2(n_2)}{2}$$

In the same way, if $n \in [n_1, \tilde{n_1}]$, we have

$$n_0 \simeq \frac{n + {H_2}^{-1} \circ H_1(n)}{2}$$

It follows that I_1 is actually composed by $\psi_1 = \frac{{H_1}^{-1}+{H_2}^{-1}}{2} \circ H_1$ and I_2 by $\psi_2 = \frac{{H_1}^{-1}+{H_2}^{-1}}{2} \circ H_2$. Thus, the intermediate histogram constructed this way is :

$$H = H_1 \circ {\psi_1}^{-1} = H_2 \circ {\psi_2}^{-1} = (\frac{{H_1}^{-1} + {H_2}^{-1}}{2})^{-1}$$

Their algorithm is a very efficient way to perform this transformation on discrete data.

3 Square Iterative Root. Theory.

Let us see how the method deriving from Axioms 1 & 2 can be performed in practical terms. Any solution ψ of the functional equation

$$\psi^{(\circ N)} = \varphi$$

is called an *iterative root* of the given function φ. Iterative roots are studied in *Iterative functional equations* by M.Kuczma, B.Choczewski and R.Ger [3]. Some of the results of this section (propositions 3, 6 and 7) can be found in this book. We reformulate them here in the image processing numerical context, which is new.

In our case, we are just interested in square iterative roots ($N = 2$) of functions from an interval into itself. We shall even limit ourselves to the case where $\varphi : [0,1] \to [0,1]$ is an increasing homeomorphism [3]. Consequently, we call *square iterative root* (s.i.r.) of φ every increasing homeomorphism $\psi : [0,1] \to [0,1]$ such that $\psi \circ \psi = \varphi$. For example, $x \to x^{\sqrt{2}}$ is a square iterative root of $x \to x^2$.

This definition leads to several questions, as the existence or uniqueness of such an iterative root. Let us start with some remarks.

- We already know that $\varphi(0) = \psi(0) = 0$ and $\varphi(1) = \psi(1) = 1$. If φ and ψ are C^1 functions, then $\varphi'(0) = \psi'(0)^2$ and $\varphi'(1) = \psi'(1)^2$.
- If $x_o \in]0,1[$ is such that $\varphi(x_o) = x_o$, then $\psi_{[0,x_o]}$ (respectively $\psi_{]x_o,1]}$) just relies on $\varphi_{[0,x_o]}$ (resp. $\varphi_{]x_o,1]}$). Then, we can limit the search for a s.i.r. to the case $\varphi(x) > x$ for all x in $]0,1[$. The case $\varphi(x) < x$ follows immediately : ψ is a s.i.r of φ iff $1-\psi(1-x)$ is a s.i.r. of $1-\varphi(1-x)$.
- If $\varphi(x) > x$ on $]0,1[$, and if ψ is a s.i.r. of φ, then $x < \psi(x) < \varphi(x)\ \forall x \in]0,1[$.

3.1 Construction of a Square Iterative Root

Proposition 3 *Let (x_0, y_0) be a point strictly located between the graph of φ and the line $y = x$. Then φ has an infinite amount of square iterative roots ψ such that $\psi(x_0) = y_0$.*

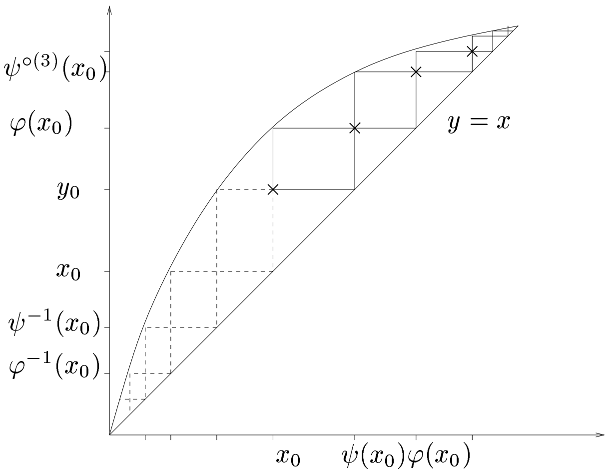

Figure 2: Construction of an Square Iterative Root.

Proof :
Let (x_0, y_0) be a point in the area described above. We want to build a square iterative root ψ of φ, such that $\psi(x_0) = y_0$. We are already able to know the sequences of positive and negative iterates of x_0 by ψ. Indeed, we know φ and

$$\forall n \in \mathbb{Z},\ \psi^{\circ(2n)}(x_0) = \varphi^{\circ n}(x_0),\ \psi^{\circ(2n+1)}(x_0) = \varphi^{\circ n}(y_0)$$

Next, since we want ψ to be an increasing function, its graph must remain in the sequence of rectangles defined by those iterations (Fig.2). Finally, if we define ψ as a continuous and increasing function on $[x_0, \psi(x_0)]$, then ψ is defined and continuous on $[0, 1]$. Indeed, ψ^{-1} is defined and continuous on $[\psi(x_0), \psi^{\circ(2)}(x_0)]$ (because ψ is an homeomorphism from $[x_0, \psi(x_0)]$ onto $[\psi(x_0), \psi^{\circ(2)}(x_0)]$). Therefore $\psi = \varphi \circ \psi^{-1}$ is also defined and continuous on $[\psi(x_0), \psi^{\circ(2)}(x_0)]$. The continuity of ψ at $\psi(x_0)$ is ensured by the construction. This process implies that the function ψ constructed is continuous on $]0, 1[$, and by extension, on $[0, 1]$.

□

In order to illustrate the process described above, we have used it to construct some square iterative roots of the function $x \to \sqrt{x}$ in Fig.3. For each choice of the point (x_0, y_0), ψ has been defined as an affine function on the set $[x_0, y_0]$.

Our problem is now to single out a root among all possible square roots, which would be more judicious than the other ones. In what follows, we analyze the properties of all square roots and discuss criteria for a best choice.

3.2 Properties of Square Iterative Roots

3.2.1 Oscillations

Proposition 4 *Let $\varphi : [0, 1] \to [0, 1]$ be an increasing homeomorphism, and ψ_1 and ψ_2 be two square iterative roots of φ. Then, for every positive number ε, there is an infinite amount of $x \in [0, \varepsilon]$ (likewise in $[1 - \varepsilon, 1]$) such that ψ_1 and ψ_2 cross each other at x.*

Proof : Assume that there exists a positive number ε such that

$$\forall x \in]0, \varepsilon[,\ \psi_1(x) \geq \psi_2(x)$$

[2] this is not specified in their paper, but it can be deduced from their code.

[3] The case "$\Phi : [0, a] \to [0, a]$" with $a > 0$ can be deduced from the previous one via the function $\varphi(x) = \frac{1}{a}\Phi(ax)$. If ψ is a s.i.r. of φ, then $x \to a\psi(\frac{x}{a})$ is a s.i.r. of Φ.

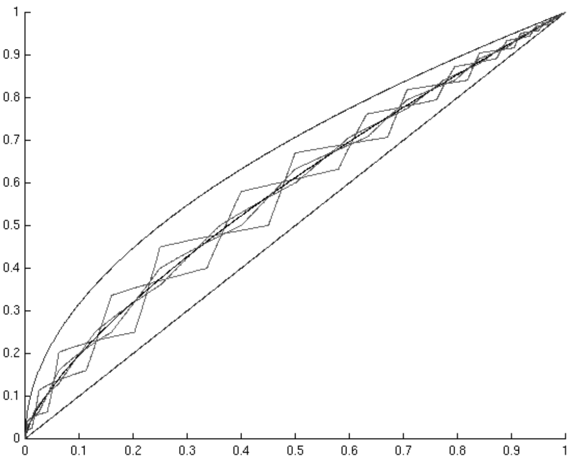

Figure 3: Some square iterative roots of the function $x \to x^{\frac{1}{2}}$. The darker curve is the explicit s.i.r. $x \to x^{\sqrt{(\frac{1}{2})}}$.

Since ψ_1 is continuous and $\psi_1(0) = 0$, there exists a positive number δ such that $\forall x \in]0, \delta[,\ \psi_1(x) \in]0, \varepsilon[$. Thus,

$$\forall x \in]0, \inf(\delta, \varepsilon)[,\ \psi_1 \circ \psi_1(x) \geq \psi_2 \circ \psi_1(x) \geq \psi_2 \circ \psi_2(x)$$

(the second inequality comes from the fact that ψ_2 is increasing). Now, the extremal terms must be equal to $\varphi(x)$. Then, all the terms are equal, ψ_1 and ψ_2 take the same values on $[0, \inf(\delta, \varepsilon)]$ and consequently ψ_1 and ψ_2 are equal everywhere (see the explicit construction).

□

3.2.2 Self-similarity

Proposition 5 *Let φ be linear on $[0, \varepsilon]$ ($\varepsilon > 0$), with a slope $a > 1$. Then, if ψ, a square root of φ, is C^1 on a neighborhood of 0, $\psi(x) = \sqrt{a}x$ on $[0, \varepsilon]$.*

Proof : Let ψ be a square iterative root of φ and x_0 be a point in $[0, \varepsilon]$. We define the sequences (x_n), (y_n) and (r_n) as $(x_n) = \psi^{-n}(x_0)$, $(y_n) = \psi(x_n)$, and

$$r_n = \frac{y_n - y_{n+1}}{x_n - x_{n+1}}$$

Then, the sequences r_{2n} and r_{2n+1} remain constant : odd rectangles (resp. even) are similar (Fig.4). Moreover, the graph of ψ in two odd (resp. two even) rectangles is the same, up to a similarity. The points (x_{2n}, y_{2n}) (resp. the points (x_{2n+1}, y_{2n+1})) are all on a line going through $(0, 0)$. Consequently, ψ is a C^1 function in a neighborhood of 0 if and only if both lines are the same for all x_0 in $[0, \varepsilon]$, i.e. iff $\psi(x) = \sqrt{a}x$ on $[0, \varepsilon]$. We can also add, following the same argument, that this function is the only square root which is concave near 0.

□

3.3 C^1 Square Iterative Roots

We focus our attention on the case when φ is a C^1 function on the open set $]0, 1[$, and when φ' doesn't reach 0..

Proposition 6 *Let φ be a C^1 increasing homeomorphism on $]0, 1[$, such that $\varphi' > 0$. There exists an infinite amount of square iterative roots ψ, which are C^1 functions on the open set $]0, 1[$. Moreover, in the previous construction, one only has to build ψ as a C^1 function on $]x_0, y_0[$ such that $\psi'(y_0^-) = \frac{\varphi'(x_0)}{\psi'(x_0^+)}$ to be sure that ψ is a C^1 function on $]0, 1[$.*

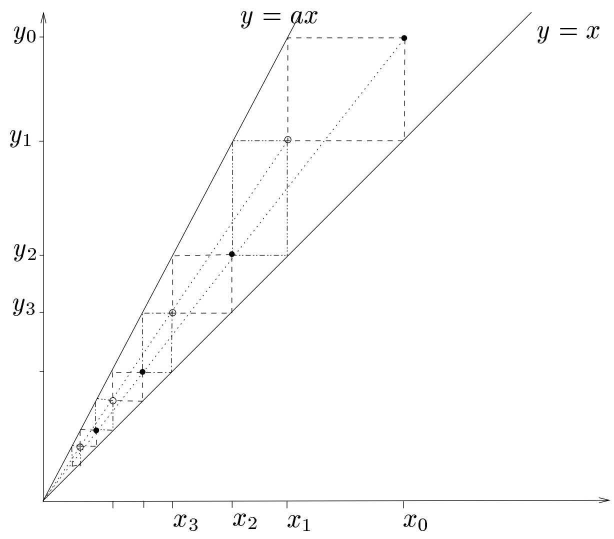

Figure 4: Behaviour of a square iterative root in the linear case.

Proof : The derivability and continuity of the derivative on the open sets are easy to prove. Let see what happens at the edges of the intervals. Let ψ be a square iterative root of φ going through (x_0, y_0). Then,

$$\forall x \in]y_0, \varphi(x_0)],\ \exists y \in]x_0, y_0],\ x = \psi(y)$$

So, if $y_0 < x \leq \varphi(x_0)$:

$$\frac{\psi(x) - \psi(y_0)}{x - y_0} = \frac{\varphi(y) - \varphi(x_0)}{y - x_0} \frac{y - x_0}{\psi(y) - \psi(x_0)}$$

ψ^{-1} is continuous, then $y \to x_0$ when $\psi(y) \to \psi(x_0)$. Thus

$$lim_{x \overset{>}{\to} y_0} \frac{\psi(x) - \psi(y_0)}{x - y_0} = \frac{\varphi'(x_0)}{\psi'(x_0{+})}$$

□

Proposition 7 *There exists a C^1 increasing homeomorphism $\varphi : [0,1] \to [0,1]$, none of whose square iterative roots is C^1 on $[0,1]$.*

Proof : We construct (fig.5) a C^1 increasing function φ, satisfying :

$$\varphi(x) = \frac{2}{3}x \text{ on } [0, \frac{1}{3}],\ \varphi(x) = 1 + \frac{1}{3}(x-1) \text{ on } [\frac{\sqrt{2}}{3}, 1]$$

Let ψ be a C^1 s.i.r. of φ. Then, ψ must at least be linear on $[0, \frac{1}{3}]$ and $[\frac{\sqrt{2}}{3}, 1]$, with respective slopes$\sqrt{\frac{2}{3}}$ and $\sqrt{\frac{1}{3}}$. Hence, $\psi(\frac{\sqrt{2}}{3}) = 1 + \frac{\sqrt{2}-3}{3\sqrt{3}}$. However, one iteration of the point $(\frac{1}{3}, \frac{\sqrt{2}}{3})$ give $\psi(\frac{\sqrt{2}}{3}) = \frac{2}{3}$. It follows that there is no C^1 function ψ, which is a s.i.r. of φ. □

The previous proposition implies that it is useless to search for some C^1 square iterative root. In what follows, we will see that the concavity of φ doesn't imply the existence of a concave square root. However, it leads us to an interesting property, which will help us to determine what kind of reasonable criterion we can use, in order to choose among all these solutions.

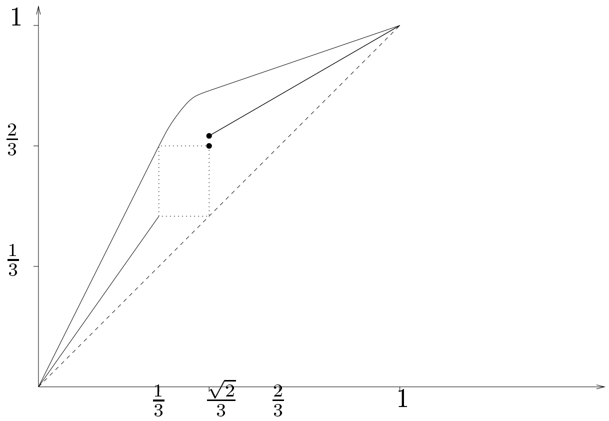

Figure 5: Example of C^1 function φ without any C^1 square iterative root.

3.4 Behavior towards the concavity of φ

The previous example is also suitable for showing that there isn't always a concave square root if φ is a concave function. Indeed, in the previous construction, we can ensure the concavity of φ. Next, if ψ is a concave s.i.r. of φ, ψ has to be linear on $[0, \frac{1}{3}]$ and $[\frac{\sqrt{2}}{3}, 1]$ (see the part about self-similarity). Now, we have :

Proposition 8 *If φ is concave near 0, and if there exists a concave square root ψ for φ, then ψ is unique.*

Proof : Assume that ψ_1 and ψ_2 are two distinct concave square roots of φ. Let $a > 0$ be such that φ is concave on $[0, a]$. It exists x_0 in $[0, a]$ such that $y_0 = \psi_1(x_0) > y'_0 = \psi_2(x_0)$. Let $x_k = {\psi_1}^{-k}(x_0)$ and $y_k = {\psi_1}^{-k}(y_0)$ be the iterates of x_0 and y_0 by ${\psi_1}^{-1}$, and (x'_k), (y'_k) those of x_0 and y'_0 by ${\psi_2}^{-1}$. All these values are of course in $[0, a]$. Then, for each positive integer k :

$$\frac{\varphi(y_{2k+2}) - \varphi({y'}_{2k+2})}{y_{2k+2} - {y'}_{2k+2}} \leq \frac{\varphi(x_{2k+2})}{x_{2k+2}}$$

Indeed, the first term is the slope of the chord between $(y_{2k+2}, \varphi(y_{2k+2}))$ and $(y'_{2k+2}, \varphi(y'_{2k+2}))$, and the second one is the slope of the chord between $(x_{2k+2}, \varphi(x_{2k+2}))$ and $(0, 0)$. Now, $0 < x_{2k+2} < y'_{2k+2} < y_{2k+2}$, thus the previous inequality follows from the concavity of φ. Consequently :

$$\forall k \geq 0, \ \frac{y_{2k} - {y'}_{2k}}{x_{2k}} \leq \frac{y_{2k+2} - {y'}_{2k+2}}{x_{2k+2}}$$

The sequence $(\frac{y_{2k} - {y'}_{2k}}{x_{2k}})$ is increasing, and then is always larger than its first term, which is positive. Now, if ψ_1 and ψ_2 are concave, they have right derivatives at 0, which must be equal to $\sqrt{\varphi'(0)}$. It follows that the sequences $\frac{y_{2k}}{x_{2k}}$ and $\frac{y'_{2k}}{x_{2k}}$ must have the same limit, and thus

$$\frac{y_{2k} - {y'}_{2k}}{x_{2k}} \overset{k\to\infty}{\to} 0$$

which contradicts the previous property. □

4 The Choice of the Square Root

4.1 How to choose the "best" square iterative root ?

Let I_1 and I_2 be the images, whose histograms have to be brought closer. According to the above considerations, we have to find a continuous function ψ, for which $\psi \circ \psi = {H_1}^{-1} \circ H_2$. Then,

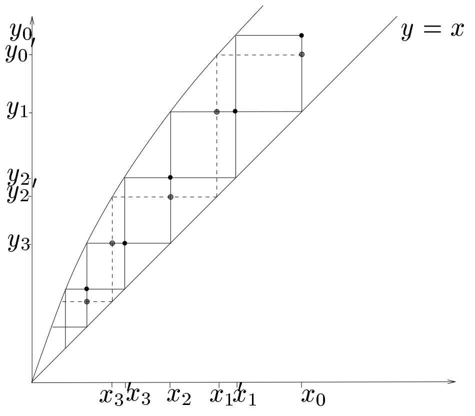

Figure 6: Iterations of ${\psi_1}^{-1}$ and ${\psi_2}^{-1}$.

$\psi(I_2)$ and $\psi^{-1}(I_1)$ will have the same histogram, which will be an intermediary between H_1 and H_2. Now, we have seen that there is an infinite amount of such functions. Even if we search C^1 square roots on the open set $]0,1[$ (in case ${H_1}^{-1} \circ H_2$ is a C^1 function itself), there exists an infinite amount of solutions. On the other hand, there exist C^1 homeomorphisms on $[0,1]$, none of whose square root is C^1 on the closed set $[0,1]$, and one can find concave functions deprived of any concave s.i.r. Thus, no usual structural criterion seems at hand to single out a right square root.

Let us come back to the previous proposition (*prop.8*): if a function φ has a concave square root ψ, then this square root is the "right" one, the one we look for. Look at Fig.3 : it shows a variety of bad solutions for $\varphi(x) = \sqrt{x}$. We have seen that those solutions are oscillating around each other, whatever φ is (*Prop.4*). In this case, each solution oscillates around the concave one $\psi(x) = x^{\frac{1}{\sqrt{2}}}$, and changes its convexity an infinite amount of times near 0 (and near 1). Thus, the concave solution is the one which "oscillates" as little as possible. In the next section, we propose a variational criterion based on a non oscillation requirement.

4.2 Behavior towards the noise

Assume that we have a uniform quantization noise b added to I ($-0.5 \leq b \leq 0.5$). Then, after a composition by a function $\psi : [0,255] \to [0,255]$:

$$\psi(I+b) \simeq \psi(I) + \psi'(I)b \ \text{ (first order)}$$

The noise becomes $\psi'(I)b$, which depends on the grey level of the point. It follows that the noise is not uniform any more. The more the variations of ψ' are important, the less the noise is uniform. Now, the uniformity of noise requirement is the most relevant one in applications. Therefore, if we want to perform a histogram modification, we must try to choose ψ' with the smallest variations around its mean. However, in our case, the contrast change used for the second image involves ψ^{-1}, we should consequently try to minimize : $\int_0^1 (\psi'-1)^2 + \int_0^1 (\frac{1}{\psi'(\psi^{-1})} - 1)^2$. Now, we can write

$$\int_0^1 (\psi'-1)^2 + \int_0^1 (\frac{1}{\psi'(\psi^{-1})} - 1)^2 = \int_0^1 (\psi'-1)^2 + \int_0^1 (\frac{1}{\psi'} - 1)^2 \psi' = \int_0^1 (\psi'-1)^2 (1 + \frac{1}{\psi'})$$

In practice, we observed that this minimization gives almost the same results as the minimization of $\int_0^1 (\psi'-1)^2$ (because ψ' is close to 1). Therefore, we shall make the approximation :
Principle : The square root which alters the least the uniformity of the quantization noise must minimize $\int_0^1 (\psi'-1)^2$.

4.3 Results on examples

We have seen and explained why the best square root for our problem is the one which "oscillates" as little as possible. On the other hand, a "good" solution, from the quantization noise point of view, has to make $\int \psi'^2$ as small as possible. In order to understand the link between both criteria, we have made the following experiment : let φ be a concave function, whose concave square root f we know. First, we compute the value y_0 in $[\frac{1}{2}, \varphi(\frac{1}{2})]$, which minimizes the discrete version of $\int \psi'^2$ (computed on the sequence of iterates of the point $(\frac{1}{2}, y_0)$). Next, we compare this value to the value of the concave s.i.r. f at $\frac{1}{2}$. The results are in the following table (we use 30 iterations of the point, and the precision on y_0 is better than 10^{-3}) :

φ	f	y_0	$f(\frac{1}{2})$
$\sqrt{x}$	$x^{\frac{1}{\sqrt{2}}}$	0.612553	0.612547
$f \circ f$	$ln((e-1)x+1)$	0.6184	0.62011
$x^{\frac{1}{4}}$	$x^{\frac{1}{2}}$	0.7074	0.7071
$f \circ f$	$\frac{4}{\Pi} \arctan x$	0.5891	0.5903

Thus, for those functions, the value y_0 computed is quite close to the value at $\frac{1}{2}$ of the solution we would like to construct.

4.4 S.I.R. Algorithm

Our algorithm is the following one (for the description, $\varphi > Id$) : first, let $x_o = \frac{1}{2}$. We choose $\psi(\frac{1}{2})$ in the open set $]\frac{1}{2}, \varphi(\frac{1}{2})[$, such that the sequence $(\psi^{\circ k}(\frac{1}{2}), \psi^{\circ(k+1)}(\frac{1}{2}))_{k\in\mathbb{Z}}$ oscillates as few as possible. Now, the values of ψ at all iterates (negative or positive) of $\frac{1}{2}$ are fixed (see Fig.2). We already know that, once $\psi(\frac{1}{2})$ is chosen, the construction of ψ as a continuous and increasing function on $[\frac{1}{2}, \psi(\frac{1}{2})]$ is enough to completely define ψ as continuous and increasing on $[0,1]$. Indeed, the values of ψ at the other points of the set $[0,1]$ will be defined as iterates of those on $[\frac{1}{2}, \psi(\frac{1}{2})]$. So, we suggest to construct ψ on a dichotomic subdivision of $[\frac{1}{2}, \psi(\frac{1}{2})]$.

The first step consists on the construction of $\psi(\frac{1}{2})$. This value must lie in $]\frac{1}{2}, \varphi(\frac{1}{2})[$, and is actually chosen in the open set such that the iterates of $(\frac{1}{2}, \psi(\frac{1}{2}))$ minimize the discrete version of ψ'^2. We follow the same rule to construct by dichotomy the values of ψ on $]\frac{1}{2}, \psi(\frac{1}{2})[$.

Then, for $k = 0..K$ (k is the current depth of the dichotomy, and K the final one), for each $(\varepsilon_0, ..., \varepsilon_k) \in \{0,1\}^k$, let

$$x_0 = (\sum_{i=0}^{k} \varepsilon_i \left(\frac{1}{2}\right)^i)(\psi(\frac{1}{2}) - \frac{1}{2}) + \frac{1}{2}$$

We want to construct $\psi(x_0)$. Now, the value chosen must remain coherent with the points previously constructed in the dichotomy. Consequently, let X_0 and X_1 be the points on both sides of x_0 which have already been dealt with in previous steps of the dichotomy :

$$X_0 = (\sum_{i=0}^{k-1} \varepsilon_i \left(\frac{1}{2}\right)^i)(\psi(\frac{1}{2}) - \frac{1}{2}) + \frac{1}{2}$$

$$X_1 = X_0 + (\frac{1}{2})^{k-1}(\psi(\frac{1}{2}) - \frac{1}{2})$$

x_0 is the middle of $[X_0, X_1]$ (see Fig.7) and $\psi(X_0)$ and $\psi(X_1)$ are already fixed. Therefore, we choose $\psi(x_0)$ in the open set $]\psi(X_0), \psi(X_1)[$, such that the sequences of the iterates of $(x_0, \psi(x_0))$ minimizes the discrete version of $\int \psi'^2$.

Thanks to this algorithm, we are able to construct a "good" square root for any function φ which doesn't cross the line $y = x$ on $]0,1[$. If a function φ is an increasing homeomorphism on $[0,1]$, we can use the same algorithm on each interval $[x_0, x_1]$ such that $\varphi(x_0) = x_0$, $\varphi(x_1) = x_1$, and $\varphi(x) \neq x$ on $]x_0, x_1[$. This way, we can compute a "good" square root for any increasing homeomorphism on $[0,1]$, and next for any increasing homeomorphism on a segment $[a,b]$.

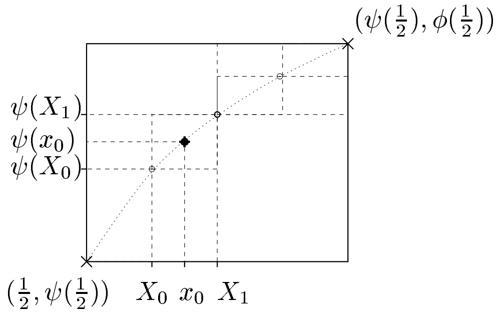

Figure 7: Construction of ψ on $[\frac{1}{2}, \psi(\frac{1}{2})]$.

5 Comparison of the methods. Experiences and applications

5.1 Theoretical Comparison of the methods

First, let us remark that if $I_2 = \lambda I_1$ then the s.i.r. equalization builds as an intermediate image between I_2 and I_1 the image $I = \sqrt{\lambda} I_1$, whereas the inverse average equalization builds $I = \frac{\lambda+1}{2} I_1$. Now, when λ is close to 1, $\sqrt{\lambda} \simeq \frac{\lambda+1}{2}$, thus the results obtained with both methods are quite similar. By the same arguments, we can show in a more general way that when ${H_1}^{-1} \circ H_2$ is close to Id, the results obtained with the two methods are almost identical.

However, the i.a.e. method is more conservative since it maintains the mean of the two images constant. Indeed, if we note $\bar{I_i}$ the mean of I_i, we have :

$$\bar{I_i} = \int_0^1 x h_i(x) dx = 1 - \int_0^1 H_i(x) dx = \int_0^1 {H_i}^{-1}(x) dx.$$

Then the average of the two means is :

$$\frac{\bar{I_1} + \bar{I_2}}{2} = \int \frac{{H_1}^{-1} + {H_2}^{-1}}{2}$$

and remains the same before and after the inverse average equalization. This is not the case for the s.i.r method : if $I_2 = \lambda I_1$, then the average of their means before the transformation is $\frac{\lambda+1}{2} \bar{I_1}$ and becomes $\sqrt{\lambda} \bar{I_1}$. This conservation principle leads us to assume that the use of the i.a.e. is certainly more relevant in iterative algorithms.

5.2 Results on a stereo pair

Fig.8 shows a pair of aerial images in the region of Toulouse. Although the angle variation between both views is small, and the photographs are taken at nearly the same time, we see that the lighting conditions vary significantly. Fig.9 shows the result of the specification of the histogram of each image on the other one. Fig.10 shows both images after equalization with the s.i.r. method, Fig.11 after inverse average equalization, and Fig.12 represents the cumulative histograms of the images before and after the different transformations. We can see that the histograms of the images after both equalizations are close and really intermediate between the previous ones. All the comparisons which follow are made between specification and the s.i.r. equalization but qualitative results would be very similar and conclusions identical with the inverse average equalization.

If we scan some image details (Fig.14 and 15), the damages caused by a direct specification become obvious. Let us specify the darker image on the brightest one. Then the information loss, due to the reduction of dynamic range, can be detected in the brightest areas. Look at the roof of the bright building in the top left corner of the image (Fig.14) : the chimneys project horizontal shadows on the roof. In the specified image, these shadows have almost completely vanished, and we can't even discern the presence of a chimney anymore. In the same image after s.i.r. equalization, the shadows are still entirely recognizable, and their size reduction remains minimal. Fig.15 illustrates the same phenomenon, observed in the bottom center of the image. The structure present at the bottom of the image has completely disappeared after specification and remains visible after midway equalization. These examples show how visual information can be lost by specification and how midway algorithms reduce significantly this loss.

5.3 Multi-Channel images

Fig.16 shows two pieces of multi-channel images of Toulouse. The first one is extracted of the blue channel, and the other from the infrared channel. Fig.17 and 18 show the same images after the s.i.r. and the i.a.e. algorithm. The main interest of matching the histograms of such images is to make their visual comparison easier and more relevant. Moreover, we observed that the results of correlation-based stereo algorihtms on multi-channel pairs were never worse and often improved after this kind of equalization.

5.4 Different shots of the same picture

Fig.20 shows two different shots of the same painting, *Le Radeau de la méduse*[4], by Théodore Géricault, found on the internet. The second one is brighter and seems to be damaged at the bottom left. Fig.21 shows the same couple after s.i.r. equalization. The results of the inverse average equalization on this example are similar.

5.5 Effects on the quantization noise

Let us see the effects of midway equalization on the quantization noise. Let I_1 and I_2 be two images, and $\tilde{I_1}$ and $\tilde{I_2}$ the same images after quantization on the integers of $[0, 255]$. We will assume as a good approximation that the quantization noise $b_i = I_i - \tilde{I_i}$ is uniformly distributed on $[-0.5, 0.5]$. Thus, $E(b_i) = 0$ and its standard deviation is $\sqrt{\int_{-0.5}^{0.5} x^2 dx} = \sqrt{\frac{1}{12}} \simeq 0.289$. The following table shows what becomes the standard deviation of b_1 and b_2 after the different transformations [5]. For example, if I_1 after specification on I_2 becomes $I_{1,2}$, and $\tilde{I_1}$ after specification on $\tilde{I_2}$ becomes $I_{\tilde{1},\tilde{2}}$, we consider that the quantization noise b_1 is transformed in $\tilde{b_1} = I_{1,2} - I_{\tilde{1},\tilde{2}}$.

	Specification	I.a.e.	S.i.r.
stereo left	1.207121	0.646560	0.442230
stereo right	0.611246	0.479748	0.538039
blue channel	1.848390	1.069263	1.040278
infrared channel	0.145176	0.961488	0.480899
meduse left	0.598254	0.458458	0.540970
meduse right	0.393870	0.354508	0.462411

We observe that the effects of the I.a.e. and of the S.i.r. on the noise are far more symmetrical than those of specification. With these methods, the standard deviation of the noise is multiplied by a number smaller than 2.5, except in the case of the multi-channel images. Indeed, those images are farther from each other and consequently subject to more drastic transformations. In this case, the ratio is smaller than 3.7.

[4]Musée du Louvre, Paris.

[5]The non-quantized images used to compute these results have been obtained in the following way : for the multi-channel images and the stereo pair, we had the original non-quantized images ; for the painting, the images were already quantized, so we added a simulated noise, uniformly distributed on $[-0.5, 0.5]$.

5.6 The Shannon point of view : effects on the spectrum

In general, the composition of an image by a function doesn't maintain the property "to be well sampled, according to Shannon-Whittaker sampling theorem". Now, we argue that when we use correlation technics on a stereo pair, in order to get the depth map of the area, the images must be well sampled if we want a good precision ([5]).

In theory, we understand that a modification of the histogram via the s.i.r. algorithm must be smoother for the spectrum than direct specification. Indeed, if $\psi : [0,1] \rightarrow [0,1]$ can be approximated by a polynomial function : $\psi(s) = s + a_2 s^2 + a_3 s^3 + ...$, then :

$$\hat{\psi(I)} = \hat{I} + a_2 \hat{I} * \hat{I} + a_3 \hat{I} * \hat{I} * \hat{I} + ...$$

The values a_i must be as small as possible in order for $\hat{\psi(I)}$ to remain well sampled. This happens if ψ remains as close to identity as possible and is smooth enough. Now, we have seen that the functions used in the s.i.r. (see for example Fig.13) and the i.a.e. methods are indeed closer to the identity than the one used in specification (${H_1}^{-1} \circ H_2$). However, it is not enough to conclude.

Consequently, in order to check experimentaly the effects of the different transformations on the spectrum, we construct the following example : the left half of each image of the couple Fig.8 is zoomed by FFT, which gives I_1 and I_2. Next, we specify I_1 on I_2 (result $\bar{I_1}$), and we compute the modification of I_1 via the s.i.r. algorithm (result $\tilde{I_1}$) or the i.a.e. (result $\breve{I_1}$). Finally, we compute the spectrums of I_1, $\bar{I_1}$, $\tilde{I_1}$ and $\breve{I_1}$ (Fig.22). We see that the spectrums of $\tilde{I_1}$ and $\breve{I_1}$ remain more in the central square than that of $\bar{I_1}$. Actually, if we compute the rate of energy which goes out of the central square after different modifications, we get 0.0107 for specification, 0.0027 for s.i.r., and 0.0013 for i.a.e. The energy stays more condensed in the spectral support of I_1 after an s.i.r. or an i.a.e. modification. Consequently, $\tilde{I_1}$ and $\breve{I_1}$ remain better sampled than $\bar{I_1}$ [6].

5.7 Comparison using the Kullback-Leibler distance

Finally, we choose to compare experimentaly the results of both midway equalizations for a statistical distance. The following table contains the Kullback-Leibler distances [7] between the histograms of the previous pairs, before, and after each kind of transformation.

	before any transformation	after I.a.e.	after S.i.r.
stereo pair	1.622677	0.005506	0.001155
Different channels	1.586096	0.008154	0.017710
Radeau de la Meduse	0.298163	0.001402	0.001435

We can observe that the algorithm giving the best performances depends on the input pair. So, these results don't allow us to conclude experimentally on the choice of the method.

6 Conclusion

This paper presented two ways of giving to a pair of images the same histogram, while remaining coherent with their previous gray level dynamics. This intermediate histogram has been constructed following two pairs of general and elementary axioms, which have led to one method or the other. We have seen that they can be applied to a wide range of image pairs, giving in each case a better representation of the images for observation, comparison or ulterior manipulations. It has been shown that those transformations are smoother for the images than direct specification

[6]Most natural images have a spectrum really concentrated in the center of its support. It follows that after a transform like a histogram modification, the spectrum sprawls but doesn't go much out of its support. The ratios of energy presented above illustrate this fact. However, we win a factor larger than 5 and this gain can be really useful in satellite imaging applications to depth map recovery.

[7]The distance between two histograms was computed symmetrically as $\frac{1}{2}(\int h_1 \log \frac{h_1}{h_2} + \int h_2 \log \frac{h_2}{h_1})$. The histograms were previously lightly smoothed in order to avoid them to vanish at some point.

: they reduce the loss of visual information generated by this kind of alteration and are better towards the quantization noise. They are also more respectful of the property "to be well sampled". The theoretical and experimental comparison of these midway equalizations was not enough to conclude on the choice of the best method. Both methods being acceptable, only practice allows to decide. From the structural point of view, the result given by i.a.e. is simpler and unique, given by an explicit formula. The s.i.r. equalization has to be considered as an alternative, and gives experimentally similar results, but its construction is certainly needlessly complicated, and more sensible to robustness problems. Consequently, we consider that the i.a.e. should be recommended as a standard method, which can be applied efficiently by the algorithm proposed by Cox et al. An extension of this theory to movie equalization is being currently studied by the author.

References

[1] I.Cox, S.Roy, S.L.Hingorani, "Dynamic Histogram Warping of Images Pairs for Constant Image Brightness", IEEE Int. Conf. on Image Processing, vol. 2, pp. 366-369, 1995.

[2] D. C. Gonzalez and R.E. Woods, *Digital Image Processing*, AddisonWesley Publishing Company, 1993.

[3] M.Kuczma, B.Choczewski and R.Ger, *Iterative Functional Equations*, Encyclopedia of mathematics and its applications, Cambridge University Press, 1990.

[4] W.Metzger, *Gesetze des Sehens*, Waldemar Kramer, 1975.

[5] B.Rougé, *Théorie de la chaîne image et restauration d'image optique à bruit final fixé*, Mémoire d'habilitation à diriger des recherches, 1997.

[6] X-D Yang, Q.Xiao and H.Raafat, "Direct mapping between histograms : An improved interactive image enhancement method", IEEE Int. Conf. on Systems, Man and Cybernetics, pp.243-247, 1991.

Figure 8: Stereo pair before equalization (aerial images of the region of Toulouse).

Figure 9: Same images after specification of their histograms on each other (left : the histogram of the first image has been specified on the second, and right : the histogram of the second image has been specified on the first).

Figure 10: Stereo pair after equalization via the s.i.r. algorithm

Figure 11: Stereo pair after equalization via the i.a.e. algorithm

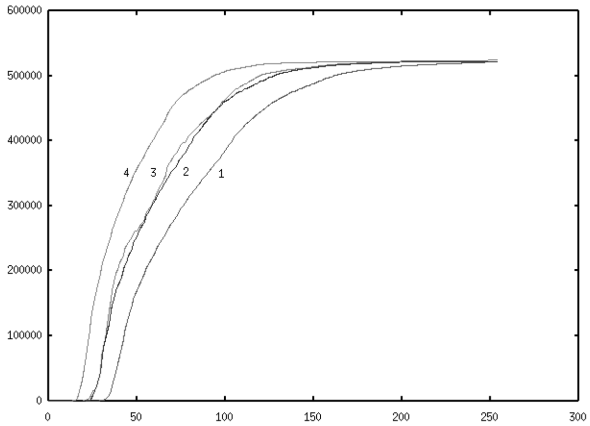

Figure 12: Cumulative histograms of the stereo pair (Fig.8) before and after each transformation. 1 and 4 : Cumulative histograms of the first and second image, 2 : Common histogram after i.a.e., and 3 : Common histogram after s.i.r. equalization

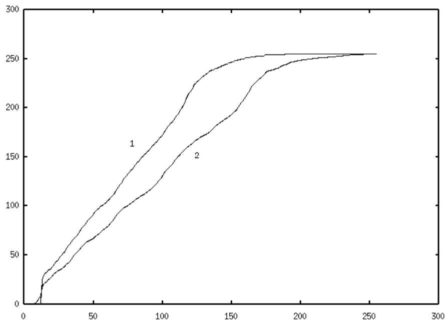

Figure 13: The first curve (1) represents ${H_1}^{-1} \circ H_2$ for the stereo pair and the second one (2) is the square iterative root computed in this case.

Figure 14: First extract of the stereo pair. From left to right : in the original image, in the specified one, in the equalized image.

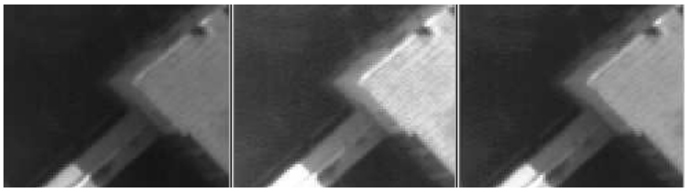

Figure 15: Second extract of the same pair. From left to right : in the original image, in the specified one, in the equalized image.

Figure 16: Images of Toulouse. Blue and infrared channel.

Figure 17: The same images of Toulouse after s.i.r. equalization.

Figure 18: Blue and infrared channel after inverse average equalization

Figure 19: Blue and infrared channel after specification on each other.

Figure 20: Two shots of the *Radeau de la Méduse*, by Géricault

Figure 21: The same images after s.i.r. equalization.

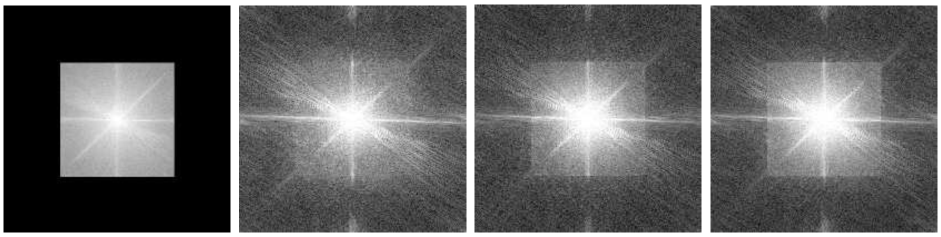

Figure 22: Spectrums of Image I_1, $\bar{I}_1$, $\tilde{I}_1$ and $\check{I}_1$. We can see that the spectrums of $\tilde{I}_1$ and $\check{I}_1$ remain more condensed in the central square than the one of $\bar{I}_1$.