Movie and Video Scale-Time Equalization Application to Flicker Reduction

Julie Delon

Abstract—Image flicker is a general film effect, which can be observed in videos as well as in old films, and consists of fast variations of the frame contrast and brightness. Reducing flicker of a sequence improves its visual quality and can be an essential first treatment before ulterior manipulations. This paper presents an axiomatic analysis of the problem, which leads to a global and fast method of "de-flicker," based on the scale-space theory. The stability of this process, called *scale-time equalization*, is ensured by the scale-time framework. Results on different sequences are given and show great visual improvement.

Index Terms—Equalization, flicker, histogram, movie, scale-space, scale-time.

I. INTRODUCTION

IMAGE flicker is a general film effect, especially known for its presence in old movies. It consists of fast variations of the frame contrast and brightness. Old movies suffer, at the same time, from age-related degradations and from the primitive technology used at the early age of cinema. In their case, flicker can be caused by physical degradations of the film, by the use of unstable chemical products, or by aberrations in the exposure time for each frame [1]. Intensity flicker can also be observed in many amateur videos, whose luminosity has not been controlled during the shooting, or more strongly in the frequent case of low time sampling, as in video surveillance.

There have been few attempts to remove flicker, probably because the problem seems quite easy to solve. However, this is not the case, as flicker can have many aspects, and is not well removed by simple affine transformations on the intensity of each frame. Most of the time, the problem is considered as a global degradation of the image. In a sense, this strong assumption is a simplification of the problem, but global methods (dependent only on the images histograms) are *a priori* much more robust to shaking, to motion and to noise presence, than local ones. After a state-of-the-art and a discussion on the subject, we will make this basic assumption in the paper. We will then see what kind of "time smoothing" generic method can be considered. Such a method should involve a scale of correction, representing the size of the time neighborhood used to change each image. This size should correspond to the limit between the flicker variations and the "natural" intensity variations in the movie. Except for this time scale, which is obviously user dependent, we look

for a method without any *a priori* assumption or user parameters.

The role of time scale, however, must be formalized and leads us to involve *scale-space theory*. This theory, founded by Witkin [2], Marr [3], and Koenderink [4], and among others pursued by Florack, Romeny, Viergever, and Lindeberg [5]–[7], links, under basic assumptions, iterated filtering to partial differential equations. It leans on the central concept of causality, which must be understood here in the sense stated by Guichard and Morel in [8]: It gathers such requirements as regularity, recursivity (semi-group structure), and a local comparison principle. As we shall see, these principles, associated with an axiomatic analysis of key examples, lead directly to a simple solution of filtering, connected with the midway image equalization presented in [9] and based on [10]. We will see why this scale-time dimension ensures the stability of the method. The end of the paper will be devoted to experiments.

II. FIRST STEPS ON THE SUBJECT

A. Some Notations

Let us begin with some notations that we will use all along the paper. Let $v : \Omega \to [0, 1]$ ($\Omega \subseteq \mathbb{R}^2$) be an image[1] and h the distribution of its intensity values (also called histogram of u). The *cumulative histogram* of v 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, $\varphi(v)$ is a new image, whose cumulative histogram is $H \circ \varphi^{-1}$.

We point out that such a transformation, called contrast change, preserves the order of the grey levels in v and does not create or cancel any image information.

We define a movie u as a sequence of images $(u_t)_{t \in \mathbb{R}}$ having the same set of definition Ω. $(H_t)_{t \in \mathbb{R}}$ denotes the sequence of its cumulative histograms. In what follows, we will always note T a "de-flicker" operator, and $\tilde{u} = T(u)$ the deflickered sequence. For a sake of simplicity, we assume that the set of definition Ω is fixed and that all the films that we consider are defined on Ω. As we will see that the results found for T are independent from Ω, this hypothesis is not restrictive.

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

B. Previous Work

The papers about flicker can be classified in two categories: those which treat image flicker as a phenomenon affecting the whole image in an equal way, and those which treat it as a local degradation. In [11] and [12], Richardson and Suter observe that a standard compression algorithm (MPEG) is not really suitable for old movies because of their obvious degradations, and propose to restore a sequence prior to compression by a motion-corrected temporal filtering technique. Before this filtering, the flicker is reduced by applying an affine stretching to the histogram of each frame, so that it fills a given dynamic range. If this global correction is simple and seems judicious at first, it is emphasized in [13] that it is not really satisfying. In [13], Decencière proposes methods to correct intensity flicker in movies, which take the basic assumption that each image suffers an affine degradation of the form

$$u_t = \alpha_t u_t^o + \beta_t$$

where u_t is the observed frame at time t, u_t^o is the original frame. If the frame at time $t-1$ has been restored, he restores the frame at time t with the assumption that some moments of the current frame remain constant (the mean and/or the extrema of the frame for instance). All the frames are restored this way. Finally, according to his conclusions, the most satisfying results are obtained by fixing the mean and the difference between the extrema of the frame. The affine correction methods have the advantage of being fast and simple, but such iterative methods involve an obvious drawback: They are always very sensitive to initialization. Here, the first frame is not restored and must be chosen very carefully, because its characteristics will be propagated to all frames. Recursive methods also lack of symmetry properties: It would be natural for a deflickering process to depend neither on the time direction, nor on the time origin. Another important drawback is the fact that there is no reason for the observed intensity degradations to be well modeled by an affine transformation.

In order to allow completely general transformations, the authors of [14] propose a solution based on histogram matching: "The purpose is to change each image histogram into a target histogram calculated as average of image histogram and its neighbors." This method is judicious and seems to give good results (an example is available at http://gpis-erver.dcom.upv.es/restoration.html). However, we will see later its drawbacks and why the scale-time equalization (STE), which has large similarities with it, is more suitable.

A first local treatment of flicker is proposed by Roosmalen *et al.* in [15]. The idea is to split the frame in overlapping blocks, and to consider that the flicker is affine on these blocks. After the local estimation of parameters, a rejection criterion is introduced in order to avoid problems due to local motion or blotches in the correction. However, this method is still sensible to outliers, and suffers from its recursive form, which creates a bias in the correction and remains very sensible to initialization. Ohuchi *et al.* [16] propose a more robust method that models the transformation between two consecutive frames by

$$u_{t+1}(x) = \alpha_t(x)u_t(x) + \beta_t(x) + \gamma_t(x)$$
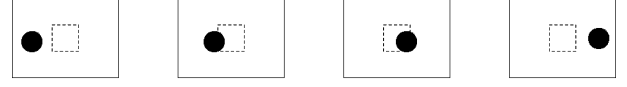

Fig. 1.   Successive frames of a moving object on a white background. The local brightness of the image changes, whereas the global histogram remains constant.

where α_t and β_t are second-order polynomials. The coefficients of α_t and β_t are estimated by minimizing a function of γ_t, chosen in order to be robust to outliers. A scale parameter decides if a local change is due to flicker or motion. This method still suffers of its lack of symmetry in the time direction. Moreover, the choice of the scale parameter is critic, and can lead to confuse a severe flicker with motion. Finally, this method relies on a specific (polynomial) model, that might not be well adapted to all examples (see Fig. 4).

C. Flicker: Local or Global? Brightness Axiom

Before trying to treat or remove intensity flicker from a sequence, we should wonder how to define it. We referred to it as variations of the frame brightness in a movie, but did not specify if those variations could be local in space or should be considered as global. As we said, papers which treat it as a local phenomenon have to involve motion detectors. Indeed, if we consider the very simple case of a dark object moving on a light background, as represented in Fig. 1, the local histogram of the central region (delimited by a doted line on the figure) changes obviously in time, whereas the global image histogram remains identical. In the same way, any motion of an object on a background could be interpreted as local flicker. Consequently, a local treatment of flicker has to involve motion detectors in order to avoid correcting local changes due to motion and not to flicker. Ultimately, even a film composed of images of a centered gaussian noise of standard deviation σ is flickering at the scale of the pixel, whereas the global brightness impression does not change. In addition to motion, a local definition of flicker is also much more sensible to global shaking (usual in old films), local blotches, or noise presence. Consequently, such a definition may well mix several phenomena, and depends a lot on the size of the chosen locality. In comparison, a global definition of flicker relies only on the histograms of the frames, which do not change much in case of shaking, object motion or in presence of jitter noise. These are principally the reasons why we chose to give up the use of the spatial and geometrical coherence between the frames, and decided to define flicker as a global phenomenon. As a consequence, if $T : u \to \tilde{u}$ is a flicker (or, in a symmetric way, deflicker) operator, we consider that

Axiom 0, Globality:

$$\forall t \in \mathbb{R}, \ \ \forall (x,y) \in \Omega, \ \ (u_t(x) = u_t(y) \Rightarrow \tilde{u}_t(x) = \tilde{u}_t(y))$$
$$\forall t \in \mathbb{R}, \ \ \forall (x,y) \in \Omega, \ \ (u_t(x) < u_t(y) \Rightarrow \tilde{u}_t(x) < \tilde{u}_t(y)).$$

This means that T affects each frame of the film by a contrast change. If the original film is $(u_s)_{s \in \mathbb{R}}$, the flicker changes it to $(\varphi_s(u_s))_{s \in \mathbb{R}}$, where each φ_s is a contrast change. As we define it, deflickering is then mostly independent from denoising and from motion detectors.
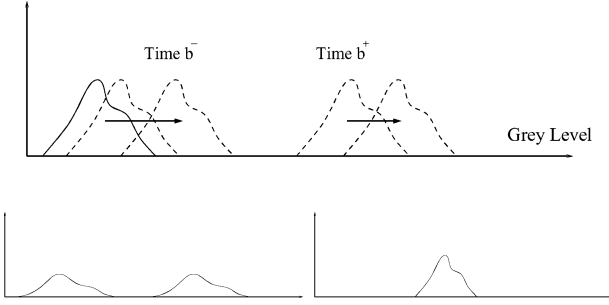
Fig. 2. Top: Successive histograms of the frames of a film, with a jump at time b. In this case, a direct weighted mean of the sequence of histograms is not adapted, in so far as it creates a new histogram with (bottom left) two modes, and, therefore, stands a chance of making some new structure appear in the film. Bottom right: Reasonable intermediate histogram.

D. Remarks on Naranjo's Method

As precised in the state of the art, affine methods [13], besides their sensitivity to initialization, are clearly not sufficient in general to model all the transformations imposed by flicker, even if we suppose that these transformations are global. The deflickering method proposed by Naranjo and Albiol in [14] is based on contrast changes: The histogram of a frame u_t is specified on a joint histogram, computed as an average of the histograms of the images u_s, $s \in [t - T, t + T]$. This can be a good and fast method in a number of cases. However, we have seen in [9] the reason why a direct average between two images histograms was not at all a satisfying intermediary between them. Let us generalize this property to the case of image sequences. We have observed in many films that the evolution of the current frame's histogram during time is not always smooth and can present severe jumps. Assume that u is composed of frames all translated from a single image u_0 : $\forall t \in [a, c]$, $u_t = u_0 + \lambda(t)$, $\lambda(t)$ having a discontinuity at b, as proposed in Fig. 2.

For a frame u_t with t near to b, a direct weighted mean of the histograms gives a result which looks like the bottom left histogram of the figure. We see that it stands a chance of creating a new histogram with several modes when all the histograms of the original frames were unimodal. The grey levels of the frame, which were concentrated around some level n_t, are separated in two distant regions, and this separation can make some new structure appear. In this case, it would be much more natural an respectful of the frame dynamics to define the new histogram as the translation of the previous ones, located around an average value, as the one represented on the bottom right of the figure. The STE that we will define in next section, relies on the same principle as Naranjo's method: replacing each image histogram by a target one, defined as a weighted intermediary between it and its neighbors. However, as we have shown (see [9] for details), the definition of an intermediary between several histograms must be understood in a "transport" sense, and not in an "average" sense.

III. SCALE-TIME EQUALIZATION

A. Action on Level Sets—Morphologic Axiom

As stated before, we impose to our operator T to follow axiom 0, i.e., to act on each image only through a contrast change.

Fig. 3. Successive frames of a film, representing a dark moving object on a flickering light background. In this kind of film, a deflickering operator should act independently on the object and on the background. This is the aim of axiom 1.

It follows that T leaves the topographic map[2] of each u_t unchanged, which naturally raises the question of the action of T on level sets.

Let us start with a very simple example, illustrated by Fig. 3. Let u be a film representing a flickering dark moving object on a flickering lighter background. We argue that, in this case, the flicker that affects the object should be treated independently of the flicker that affects the background. This strong assumption compensates the globality imposed by axiom 0: It localizes the action of the operator T separately on the object and on the background (see also the first example in Section IV, and Fig. 4). In this example, the object corresponds to the p darkest pixels of the image, and the background to the $N - p$ lightest ones. If we note, for every λ in $[0, 1]$

$$\chi^\lambda(t) = \{x \in \Omega, H_t(u_t(x)) \le \lambda\}$$
$$\chi_\lambda(t) = \{x \in \Omega, H_t(u_t(x)) > \lambda\}$$

then $\chi^\lambda(t)$ corresponds to the $\lambda\%$ darkest pixels of u_t, and $\chi_\lambda(t)$ to the $(1 - \lambda)\%$ lightest ones. Thus, the previous statement can be translated by the following axiom.

Axiom 1, Figure/Background Independence: For every λ in $[0, 1]$, the action of T on the sequence $t \to \chi^\lambda(t)$ is independent of the action of T on the sequence $t \to \chi_\lambda(t)$.

If this property holds, it is equivalent to say that T acts independently on each "level set of rank λ" sequence.

This axiom simplifies considerably the possibilities of action for T. Indeed, let us introduce the set $\mathcal{C}$ of films defined on Ω and constituted of constant images (i.e., each image of the film is a constant, but the value of the constant changes with time). We know from axiom 0 that $\mathcal{C}$ is stable under the action of T. Moreover, there is an obvious bijection between $\mathcal{C}$ and $\mathcal{F}_\mathbb{R}$, the set of real functions. Thus, the action of T on $\mathcal{C}$ can be completely represented by the action of an operator $\breve{T}$ on $\mathcal{F}_\mathbb{R}$. Now, we have the following result.

Proposition 2: The action of an operator T satisfying axioms 0 and 1 only depends on its action on $\mathcal{C}$, the set of films constituted of constant images. More precisely, if $\breve{T}$ denotes the corresponding action on the set of real functions, the action of T on any film u consists in separate actions of the operator $\breve{T}$ on the functions $t \to H_t^{-1}(\lambda)$.

Each function $t \to H_t^{-1}(\lambda)$ gives the evolution during the time of the grey levels of rank λ in the images of the film. The study of the action of T on films constituted of constant images is the aim of one of the following sections.

---

[2]The topographic map of an image v is the set of its level sets, ordered by inclusion and taken independently from their grey levels. If φ is a strictly increasing function, $\varphi(v)$ has the same topographic map as v (see [17] for more information on topographic maps).
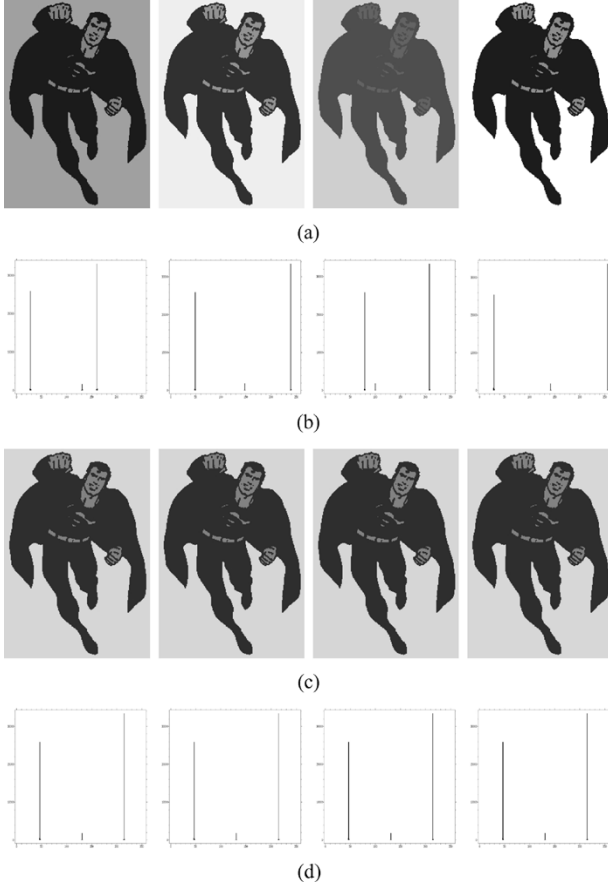
Fig. 4. (a) Successive images with the same geometry, where flicker has been added independently to the 3 different level sets. (b) Corresponding histograms. We see that there is no affine transform that maps one histogram on another. (c) Same images after STE at a large scale. The values of the level sets have been regularized in time, and the geometry does not change; hence, the four images become identical. (d) Resulting histograms of the four previous images. As expected, they are identical.

B. Scale-Time Structure of Deflicker

Before studying the particular case of a film of constant images, there is still a general and simple requirement that we can impose on T. The kind of transformation that we want to perform should naturally not depend on the choice of the time origin for the film, and should also give the same result if we reverse the time direction.

Axiom 2, Time Shift Invariance: T must commute with translations and symmetries in time

$$\forall u, \ \forall \mu \in \mathbb{R}, \ T\left(\tau_\mu(u)\right) = \tau_\mu\left(T(u)\right)$$
$$\forall u, \ T\left((u_{-t})\right) = (T(u)_{-t}).$$

This axiom leads actually to the question of the link between the operator T and the time: If the intensity flicker has to be removed, we must avoid to make disappear the natural variations of luminosity in the film, which are much slower. This need becomes more and more necessary when the length of the sequence increases. Hence, the reduction of flicker should be seen as a scale-time problem. A flicker reduction method should be flexible enough to allow the choice of a "scale" of correction, corresponding to the limit between intensity variations due to flicker, and those "natural" to the movie.

A natural way to keep the possibility of luminosity variations from one image to another after the treatment, is to use, instead of T, a family $(T_s)_{s \in \mathbb{R}}$ of deflicker operators, the scale parameter s being linked with the depth of correction. Each element of the family is obviously supposed to follow the axioms 0, 1, and 2, and the next section should allow us to determine the form of the operators T_s.

C. Particular Case of a Film Constituted of Constant Images

In this section, we will end up with the formulation of axioms for deflicker. Surprisingly (as shown previously), it is enough to state them in the case of constant images movies. This very intuitive framework will lead to proposition 3.

Now, let $c = (c_t)_{t \in \mathbb{R}} \in \mathcal{C}$, the set of films on Ω constituted of constant images. We study here the action on c of a family $(T_s)_{s \in \mathbb{R}}$ of operators following axioms 0 to 2. First, if all of the constants c_t are identical, equal to the constant μ, we impose naturally that each of the T_s keeps the film unchanged.

Axiom 3, Preservation of Constants: If $\exists \mu, \forall t \in \mathbb{R}, c_t = \mu$, then $\forall s, T_s(c) = c$.

The action of the operators $(T_s)_{s \in \mathbb{R}}$ on a sequence $\{c(t), t \in \mathbb{R}\}$ is equivalent to a smoothing process in one dimension. By making once more the parallel between $\mathcal{C}$ and $\mathcal{F}_{\mathbb{R}}$, **we call** $(\breve{T}_s)_{s \in \mathbb{R}}$ **the corresponding smoothing process on** $\mathcal{F}_{\mathbb{R}}$, **in order not to mistake its action for that of** $(T_s)_s$. The results found for $(\breve{T}_s)$ will determine the action of (T_s) on films via its separated actions on the "rank" sequences $(H_t^{-1}(\lambda))_t$.

At this time of the reasoning, any family of one dimension smoothing operators following axioms 2 (time shift invariance) and 3 (preservation of the constants) could be considered as convenient for $(\breve{T}_s)$. However, some natural assumptions, as the classical assumptions of scale-space theory, can be made on this process. Several axiomatics are possible (see [5] for a review of the main ones), depending on the authors, but they all lead to the same conclusion: The convolution with a gaussian kernel. We choose here the most general list of axioms, as stated in [8]. First, we impose a pyramidal structure on the family, which means that the signal at scale t can be computed from the signal at scale $t - h$, for h small enough. Next, we state that the family $(\breve{T}_s)$ should naturally satisfy the following local comparison principle: If c is locally larger than c', it remains larger for a smoothing small enough. Finally, it seems sensible to impose a regularity assumption on the family $(\breve{T}_s)_{s \in \mathbb{R}}$. All of these assumptions compose the following axiom.

Axiom 4, Scale-Time Structure of $(\breve{T}_s)$: The scale-time family $(\breve{T}_s)_{s \in \mathbb{R}}$ is causal, which means that $(\breve{T}_s)_{s \in \mathbb{R}}$

- is pyramidal: $\breve{T}_{s+h} = \breve{T}_{s+h,s}\breve{T}_s$, $\breve{T}_0 = Id$;
- satisfies a local comparison principle: if $c(t) > c'(t)$ for t in a neighborhood of t_0 and $t \neq t_0$, then for h small enough

$$(\breve{T}_{s+h,s}c)(t_0) \geq (\breve{T}_{s+h,s}c')(t_0)$$

  and if $c \geq c'$ everywhere, then $\forall h, \breve{T}_{s+h,s}c \geq \breve{T}_{s+h,s}c'$;
- is regular: for c a quadratic form, there exists a function F continuous with respect to the first variable, such that when $h \to 0$

$$\frac{(\breve{T}_{s+h,s}c - c)(x)}{h} \to F\left(D^2c(t), Dc(t), t, c(t), s\right).$$

Finally, for the sake of simplicity, we choose to impose an axiom of linearity on $(\breve{T}_s)_{s \in \mathbb{R}}$.

Axiom 5, Linearity: The scale-time $(\breve{T}_s)_{s \in \mathbb{R}}$ is linear.

This last axiom is a deliberated choice which allows to determine completely the action of $(\breve{T}_s)_{s \in \mathbb{R}}$ (see [8]).

Theorem 1: Let $(\breve{T}_s)_{s \in \mathbb{R}}$ be a "scale-time" family of operators acting on functions $c : \mathbb{R} \to \mathbb{R}$ and satisfying the axioms 2 to 5. Then, for each function $c : \mathbb{R} \to \mathbb{R}$, and up to a rescaling $s' = f(s)$

$$\breve{T}_s c = G_s *_t c, \;\; where \;\; G_s(t) = \frac{1}{(4\pi s)^{\frac{1}{2}}} e^{-\frac{t^2}{4s}}.$$

Other solutions, nonlinear, could be considered for the action of the family. The quality of the results given by this specific process will be obviously its best justification. Let us mention that the median filter on a neighborhood might be an alternative choice, more adapted to scene changes.

D. STE

The association of proposition 2 and of the previous theorem gives the following result, where $*_t$ denotes the convolution in the time direction.

Proposition 3: Let $(T_s)_{s \in \mathbb{R}}$ be a family of operators acting on films and satisfying axioms 0 to 3. We call $(\breve{T}_s)_{s \in \mathbb{R}}$ the family acting on real functions and reproducing the action of (T_s) on $\mathcal{C}$, the set of films of constant images. We suppose that $(\breve{T}_s)_{s \in \mathbb{R}}$ satisfies axioms 4 and 5. Then, there exists a rescaling $s' = f(s)$, such that the action of (T_s) on any film u can be described via the formula

$$\breve{T}_s \left(t \to H_t^{-1}(\lambda) \right) = G_s *_t \left(t \to H_t^{-1}(\lambda) \right)$$

where

$$G_s(t) = \frac{1}{(4\pi s)^{\frac{1}{2}}} e^{-\frac{t^2}{4s}}.$$

Proof: Proposition 2 tells us that the action of each operator T_s of the family on a film u can be described by the separated actions of the corresponding operator $\breve{T}_s$ on the functions $t \to (H_t^{-1}(\lambda))_t$ (H_t being the cumulative histogram of the image u_t). Now, theorem 1 determines completely the action of the family $(\breve{T}_s)$. ∎

The previous theorem describes completely the action of a family of operators following the axioms previously stated. This action is nothing more than a convolution of each "rank function" of the film by a gaussian. It will be called STE. In order to improve the understanding of this action, we shall now study its stability and its effects on different moments of the frames. In the last section, we will experiment it on sequences presenting an artificial or a real flicker.

E. Stability

First, let us study the effect of this time-smoothing process on the mean of the frames. If we note $m_u(t)$ the mean of the image u_t, we have

$$m_u(t) = \int\limits_0^1 x h_t(x) dx = 1 - \int\limits_0^1 H_t(x) dx = \int\limits_0^1 H_t^{-1}(x) dx.$$

Consequently, any linear operation on the sequence H_t^{-1} generates the same linear operation on the function $m_u(t)$. Hence, the following proposition.

Proposition 4:

$$\forall t_0, \quad m_{T_s(u)}(t_0) = G_s *_t (m_u)(t_0).$$

The mean function is smoothed by the heat equation. In the same way, every value α_t which can be written linearly from the values $H_t^{-1}(x)$ is also smoothed by the heat equation, as the median, for example, which is actually equal to $H_t^{-1}(1/2)$. The behavior of the mean function in a real example and for different choices of s can be observed in the experiment section [Fig. 7(c)].

The notion of stability here is closely related to the concrete application of the method. Indeed, in the theoretical study, we considered the film u as if it was infinite, which is of course not the case. The movie u has a finite discrete length P. In order to apply the previous method, we consider u as a time-periodic function. There is no reason for introducing an interaction between the first and the last image of the film, so we choose to symmetrize u on $[0, 2P]$ before its periodization. Hence, u is considered as periodic of period $2P$.

The smoothing process at scale s changes each inverse cumulative histogram $\lambda \to H_{t_0}^{-1}(\lambda)$ into $\lambda \to (G_s *_t H_t^{-1}(\lambda))(t_0)$, which converges uniformly, when $s \to \infty$, toward a limit function $\lambda \to (1/2P) \int_0^{2P} H_t^{-1}(\lambda) dt$. And finally, if we define

$$H_{\lim}(\lambda) = \left(\frac{1}{2P} \int\limits_0^{2P} H_t^{-1}(\lambda) dt \right)^{-1}$$

we can show the following result of convergence.

Proposition 5: If $(t, \lambda) \to H_t^{-1}(\lambda)$ is in $L^1([0, 1] \times [0, 2P])$ and equicontinuous in λ, then, for all t, the sequence of functions $(\lambda \to H_t^s(\lambda))_s$ converges uniformly toward $H_{\lim}$ when $s \to \infty$.

This stability result supports the idea to use large scales in order to equalize a sequence. In a sense, the scale represents the size of the time neighborhood that we choose to equalize each image of the film. The larger the scale is, the more the final histograms of the sequence will be similar and tend toward $H_{\lim}$. If a sequence does not contain scene transitions, which is the case in the experiments of Section IV, we can consider as reasonable to choose $s \geq 10$.

IV. SOME EXPERIMENTS

We have tested our algorithm on several sequences, presenting either an artificial or a real flicker. The results are presented here via some images of each sequence before and after the equalization process. However, it is not easy to evaluate the quality of the results on still images, and the improvement is much more obvious seeing the films in motion: The sequences used here and their corrections by STE can be found at http://www.cmla.ens-cachan.fr/~delon. One of the advantages of the globality hypothesis relies on the speed of the process: For images of 720×576 size, the whole process treats around four images by second on a 800-MHz PC, and 18 images by second for a 240×180 size. Let us begin with two sequences presenting an artificial flicker.

A. Sequences With an Artificial Flicker

The first one is a very simple synthetic case, which tends to illustrate the works of the method. The images of the film, that can be seen on Fig. 4(a), are all made from a single image. This original image has only three different grey levels: The first one corresponds to the clothes of the character, the second one to its skin, and the last one to the background. The different images of the film are obtained by making these values vary independently in time (but without changing the global order of the values; the background is always the lighter one, etc.). This way, each image of the film has an histogram composed of three peaks, located at grey values that change in time [Fig. 4(b)]. In this example, the artificial flicker is clearly global, but completely nonaffine. It cannot either be modeled by the local model of Ohuchi's paper [16]. Moreover, if we tried to replace each histogram by a direct weighted average of its neighbor histograms (such as that suggested by Naranjo and Albiol in [14]), the result would be catastrophic: It would create histograms with several peaks distributed all over the possible range, which has no sense. Now, as we have seen, the STE compute for each level set an average in the time direction of the grey values taken by the set. Which means here that the grey values of the background will be equalized together, etc. All the images of the film here have the same geometry, therefore, at a large scale, all the images obtained are identical [Fig. 4(c)]. The corresponding limit histogram [Fig. 4(d)] is composed of the same three peaks, located at the computed average values. We can remark that if the character was moving in position from one image to the other; the result would not change, as it just relies on histograms.

The second example is a short film shot outdoor, to which we added an artificial flicker. Concretely, each image u_t is replaced by an image $\varphi_t(u_t)$, where φ_t is a polynomial function whose coefficients change with time. Fig. 5(a) shows three images of the degraded film (there is ten frames between two consecutive images), and their intensity histograms, which look pairwise very far from each other. On the same figure, we represented the corresponding restored images and their intensity histograms. We can see that the histograms of the frames after restoration are almost identical, because of the large scale chosen for restoration. The subjective quality of the result is certain: Visually, we can not distinguish this restored sequence from the original one (before the degradation).

This example gives us the possibility to study the sensibility of our algorithm to the presence of noise. Theoretically, if we add a noise b to an image u, its intensity distribution becomes $h_u * h_b$, where h_b is the noise distribution. If the noise is gaussian, the histogram h_u is smoothed by this operation, which should not be at all disturbing for the STE process. If the noise is an impulse noise, added to $p\%$ of the pixels, its distribution is of the form $h_b = (1 - p)\delta_0 + p$. Thus, $h_u * h_b = (1 - p)h_u + p\bar{h_u}$, which still should not be much disturbing for the restoration process when p is not too large. In order to test it concretely, we made the following experiment: Before the flickering process, we added a gaussian noise of standard deviation σ to each image of the original film and restored this new sequence with our algorithm. For each image u of the original sequence, u_{noise} denotes the corresponding
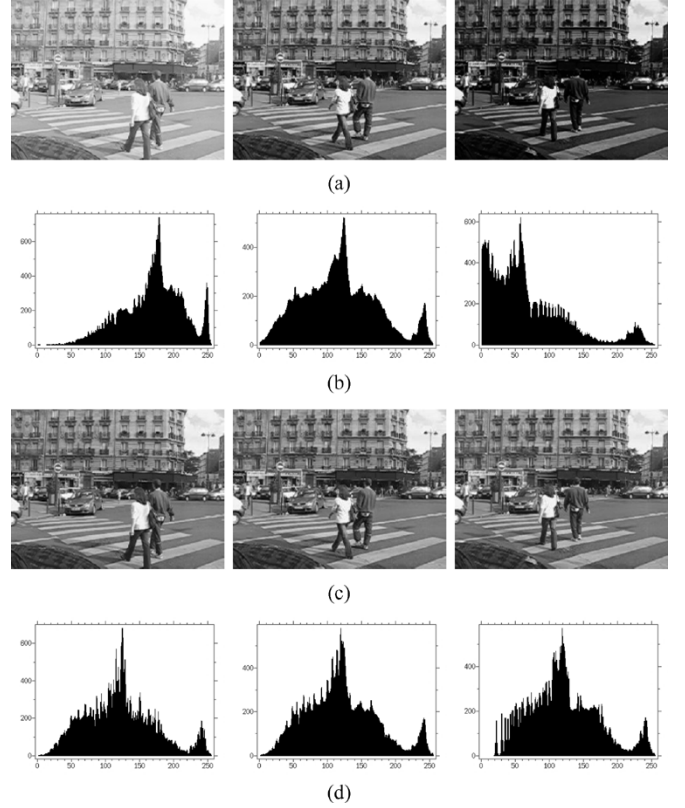

Fig. 5. (a) Three images of a film to which we added a global and highly nonaffine artificial oscillating flicker and (b) their intensity histograms. (c) Same images after scale size equalization at a large scale. The artificial flicker has completely vanished. The original sequence is restored. (d) Intensity histograms of the restored images. We see that they are very close to each other.

noised image, $u_{\text{noise+flicker}}$ the corresponding noised image after the flickering process, and u_{restored} the same image after deflickering. For each image of Fig. 5, we represented on Fig. 6 the images $u_{\text{noise+flicker}}$ and u_{restored} for $\sigma = 60$. Visually, the results are quite good, even for this large value of σ. Logically, u_{restored} should be closer to u_{noise} than $u_{\text{noise+flicker}}$. In order to verify it, we measured, for different values of σ, the average L^2 norms of the differences $u_{\text{noise}} - u_{\text{restored}}$ and $u_{\text{noise}} - u_{\text{noise+flicker}}$ for the three images of Fig. 5.

σ(noise std. dev.) :	0	10	30	60
$\|u_{\text{noise}} - u_{\text{noise+flicker}}\|_2$				
image 1	51.62	51.45	50.01	46.45
image 10	15.84	15.76	15.32	14.13
image 20	53.98	53.74	52.25	47.97
$\|u_{\text{noise}} - u_{\text{restored}}\|_2$				
image 1	5.25	5.43	5.06	4.43
image 10	3.93	4.55	4.25	4.58
image 20	4.74	5.20	5.70	8.16.

As expected, we see that the method is very robust to gaussian noise. The global L^2 difference between each image and its corresponding restored image does not change much when σ increases. We begin to see differences around $\sigma = 60$, but they remain reasonable: For $\sigma = 100$, the ratio $\|u_{\text{noise}} - u_{\text{restored}}\|_2 / \|u_{\text{noise}} - u_{\text{noise+flicker}}\|_2$ is still around $1/5$ for most of the images. The results for an impulse noise
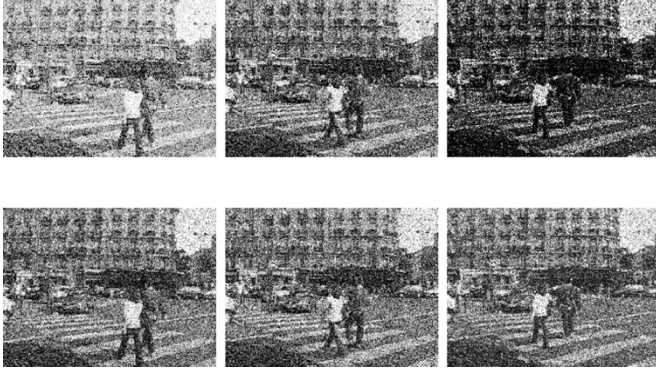
Fig. 6. Top: Images of Fig. 5, to which we added a gaussian noise of standard deviation $\sigma = 60$ before adding the flicker. Bottom: Same images after STE.

are quite similar: For a noise proportion of 30%, the previous ratio is still around 1/10 for images 1 and 20, and around 1/5 for image 10.

B. Sequences With a Real Flicker

The first real sequence is an extract of Chaplin's 1915 movie *His New Job*, and has been found on the page proposed by Naranjo and Albiol in their paper [14]. Despite the fact that the camera is fixed and the scene is shot indoor, the sequence suffers from a severe real flicker. This proves that the flicker can come from different sources, and be produced by nonnatural causes (certainly the aging of the film in this particular case). Fig. 7(a) shows three samples of the film, taken at equal intervals of time, and the same images after the STE. We can observe (even if it is more obvious seeing the film in motion) that the flicker has globally almost vanished. The same figure shows the time-evolution of the current frame mean when s increases. As predicted in *Proposition 4*, this mean function is just smoothed by the heat equation at scale s. When $s \to \infty$, the mean function converges toward the constant function computed as the average of all the frames means.

The second real example (Fig. 8) is an extract of another Chaplin movie, called *The Cure* (1917). The real flicker present in this part of the film is much worse than in the first example. In particular, some images are very dark and low contrasted. After an equalization at a large scale, we can see that the flicker has almost vanished. However, the quantization noise has globally increased, especially in low-contrasted zones. This drawback is inherent in any deflickering method, local or global, affine or not: Indeed, it appears as soon as the grey level dynamic of an image is enlarged, even locally. A good way to avoid it is to bound the slope of the contrast function that is applied to the image for restoring it.

V. Conclusion

This paper presents a global way of reducing the effects of flicker in films by a scale-time process. This process has been constructed following a group of general and elementary axioms, such as figure/background independence or time-shift invariance, associated with basic axioms of scale-space theory. As a global process, this method is robust to gaussian noise, shaking, and motion. This simple method is more generic
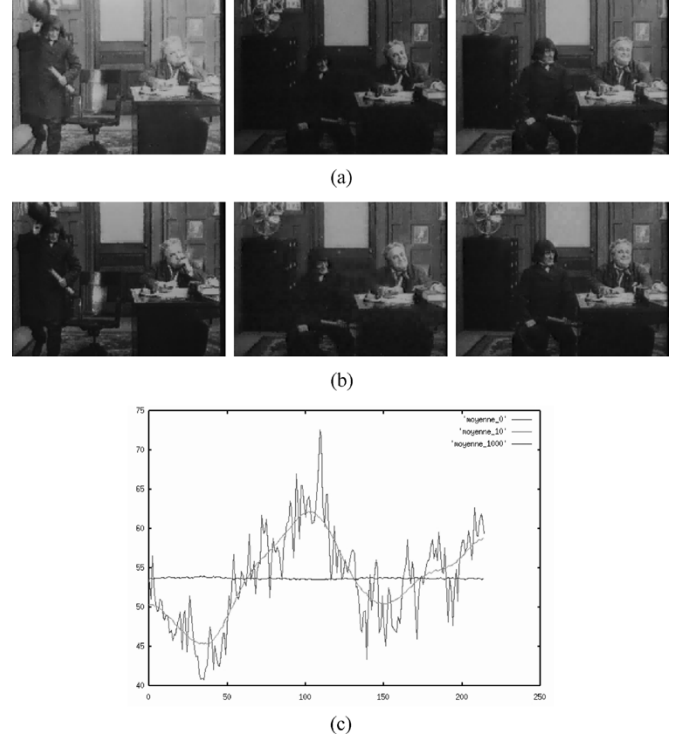

Fig. 7. (a) Three images of Chaplin's film *His New Job*, taken at equal intervals of time. This extract of the film suffers from a severe real flicker. (b) Same images after the STE at scale $s = 100$. The flicker observed before has globally decreased. (c) Evolution of the mean of the current frame in time and at three different scales. The most oscillating line is the mean of the original sequence. The second one is the mean at scale $s = 10$. The last one, almost constant, corresponds to the large scale $s = 1000$. As expected from *Proposition 4*, the mean function is smoothed by the heat equation.


Fig. 8. Top: Three images of Chaplin's film *The Cure*, taken at equal interval of time. This part of the film presents a strong real flicker. Some images in particular are very dark and low contrasted. Bottom: Same images after a STE at a large scale. The flicker has almost vanished.

than global affine methods, and we have seen why it is more respectful of the grey level dynamics of the frames than the method proposed by Naranjo and Albiol in their paper. The scale-time dimension leads to simple results of stability, ensures the robustness of the method to blotches or impulse noise, and guarantees that no bias or deviation can appear in time, contrary to recursive methods. Results on sequences containing an artificial or a real flicker are convincing: The subjective quality of the films has clearly improved. In all cases, the flicker becomes almost imperceptible. In cases of a flicker mixing both very

local changes and global oscillations, this process can still be used as a first step of deflickering before a local treatment. Now, it is clear that such a deflickering process looses its sense in the presence of scene transitions. It should, therefore, be associated with a well-chosen time-cut detector (for instance, see [18]).

REFERENCES

[1] A. Kokaram, *Motion Picture Restoration*. New York: Springer, 1998.
[2] A. P. Witkin, "Scale-space filtering," in *Proc. Int. Joint Conf. Artificial Intelligence*, 1983, pp. 1019–1021.
[3] D. Marr, *Vision*. San Francisco, CA: Freeman, 1982.
[4] J. J. Koenderink, "The structure of images," *Biol. Cybernet.*, vol. 50, no. 5, pp. 363–370, 1984.
[5] T. Lindeberg, *Scale-Space Theory in Computer Vision*. Norwell, MA: Kluwer, 1994.
[6] ——, "On the axiomatic foundations of linear scale-space: combining semi-group structure with causality vs. scale invariance," presented at the Gaussian Scale-Space Theory Conf., 1997.
[7] T. Lindeberg and B. ter Haar Romeny, "Linear scale-space," presented at the Geometry-Driven Diffusion in Computer Vision Conf., 1994.
[8] F. Guichard and J. M. Morel, *Image Analysis and p.d.e.'s*, 2002, to be published.
[9] J. Delon, "Midway image equalization," *J. Math. Imag. Vis.*, vol. 21, no. 2, pp. 119–134, Sep. 2004.
[10] I. Cox, S. Roy, and S. L. Hingorani, "Dynamic histogram warping of images pairs for constant image brightness," in *Proc. IEEE Int. Conf. Image Processing*, vol. 2, 1995, pp. 366–369.
[11] P. Richardson and D. Suter, "Restoration of historic film for digital compression: a case study," in *Proc. Int. Conf. Image Processing*, vol. 2, Washington, DC, USA, 1995, pp. 49–52.
[12] ——, "Historical film restoration and video coding," in *Proc. PCS*, 1996, pp. 389–394.
[13] E. D. Ferrandière, "Restauration automatique de films anciens," Ph.D. dissertation, Ecole Nat. Sup. des Mines de Paris, Paris, France, 1997.
[14] V. Naranjo and A. Albiol, "Flicker reduction in old films," presented at the IEEE Int. Conf. Image Processing, Vancouver, BC, Canada, Sep. 2000.
[15] P. van Roosmalen, R. L. Lagendijk, and J. Biemond, "Correction of intensity flicker in old film sequences," *IEEE Trans. Circuits Syst.*, vol. 9, no. 10, pp. 1013–1019, Oct. 1999.
[16] T. Ohuchi, T. Seto, T. Komatsu, and T. Saito, "A robust method of image flicker correction for heavily corrupted old film sequences," presented at the Int. Conf. Image Processing, Sep. 2000.
[17] V. Caselles, B. Coll, and J. M. Morel, "Topographic maps and local contrast changes in natural images," *Int. J. Comput. Vis.*, vol. 33, no. 1, pp. 5–27, 1999.
[18] V. Naranjo and A. Albiol, "Low complexity cut detection in the presence of flicker," presented at the IEEE Int. Conf. Image Processing, Vancouver, BC, Canada, Sep. 2000.
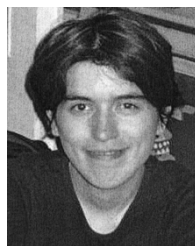
Julie Delon was born in France in 1978. From 1997 to 2001, she studied applied mathematics at the Ecole Normale Supérieure de Cachan, Cachan, France, and received the Ph.D. degree, with a dissertation on image analysis, from the CMLA, Cachan, in December 2004.

Since October 2005, she has been with Telecom Paris, Paris, France, as a CNRS researcher