APPROXIMATION ORDER OF THE LAP OPTICAL FLOW ALGORITHM

Thierry Blu[1]*, Pierre Moulin*[2]*, and Christopher Gilliam*[1]

[1]Department of Electronic Engineering, The Chinese University of Hong Kong
[2]Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign

email: {tblu,cgilliam}@ee.cuhk.edu.hk, moulin@ifp.uiuc.edu

ABSTRACT

Estimating the displacements between two images is often addressed using a small displacement assumption, which leads to what is known as the optical flow equation. We study the quality of the underlying approximation for the recently developed Local All-Pass (LAP) optical flow algorithm, which is based on another approach—displacements result from filtering. While the simplest version of LAP computes only first-order differences, we show that the order of LAP approximation is quadratic, unlike standard optical flow equation based algorithms for which this approximation is only linear. More generally, the order of approximation of the LAP algorithm is twice larger than the differentiation order involved. The key step in the derivation is the use of Padé approximants.

Index Terms— Optical flow, all-pass filtering, approximation, Padé approximant.

1. INTRODUCTION

The 2D optical flow problem consists in estimating space-varying displacement vectors $\mathrm{u}(x,y)=(u_x(x,y),u_y(x,y))^{\mathrm{T}}$ that relate two known images $I_1(\mathrm{r})$ and $I_2(\mathrm{r})$; i.e., under the ideal brightness consistency hypothesis [1]

$$I_2(\mathrm{r}) = I_1(\mathrm{r}-\mathrm{u}(\mathrm{r}))$$

where $\mathrm{r}=(x,y)^{\mathrm{T}}$ are spatial coordinates. This is a challenging problem that finds applications in a wide range of fields like computer vision, medical imaging [2, 3], biology [4, 5], and image compression. The dominant algorithms use ideas that were initially proposed in the 1980s: first, linearising the effect of small displacements to obtain the "optical flow equation". Then, using this equation as a data term in a regularization functional to be minimized (Horn-Schunck approach [6]), or as a set of constraints to be fitted blockwise using few parameters (Lucas-Kanade's approach [7]).

The type of objective function that has to be minimized in the Horn-Schunck approach has been the source of constant developments: robust penalty terms [8, 9], L_1 regularization [10, 11] and low-rank regularizers [12]. For a complete review of the state-of-the-art see [13, 14, 15, 16, 17], and, more recently, [1, 18].

In this paper, we are interested in the quality of approximation underlying optical flow algorithms. Specifically, we evaluate this quality for a new algorithm that models displacements as local all-pass (LAP) filtering operation [19]. The contribution of this paper is to analyze how the LAP algorithm makes it possible to achieve a higher order of approximation than the algorithms based on the optical flow equation, without requiring to compute higher order derivatives.

Note that this algorithm is not related to *spatio-temporal* filtering algorithms [20, 21] which rely on the time variation of the spatio-temporal Fourier phase of a *sequence of images*: only spatial filters are involved in the LAP algorithm, and it is between two images only that the displacement field is to be estimated.

2. APPROXIMATION ORDER

Usual optical flow algorithms are based on an approximation of the displacement by the vector field $\mathrm{u}(\mathrm{r})$. Using such an approximation is important in order to separate $\mathrm{u}(\mathrm{r})$ from $f(\mathrm{r})$ and so, to derive efficient algorithms. The standard approach consists in deriving an optical flow equation [6] which usually amounts to approximating $I_1(\mathrm{r}-\mathrm{u}(\mathrm{r}))$ using a first order Taylor expansion; i.e. for small values of $\mathrm{u}(\mathrm{r})$ and assuming that the image is at least twice boundedly differentiable:

$$\begin{gathered}I_1(\mathrm{r}-\mathrm{u}(\mathrm{r})) = I_1(\mathrm{r}) - \mathrm{u}(\mathrm{r})^{\mathrm{T}}\nabla I_1(\mathrm{r}) + \mathrm{O}\big(\|\mathrm{u}(\mathrm{r})\|^2\big)\\ \rightsquigarrow I_2(\mathrm{r}) \approx I_1(\mathrm{r}) - \mathrm{u}(\mathrm{r})^{\mathrm{T}}\nabla I_1(\mathrm{r})\end{gathered}$$

Here and throughout this paper, the notation $f(x)=\mathrm{O}(g(x))$ means that there exists a constant (independent of x) such that

$$|f(x)| \le \mathrm{const} \times |g(x)|.$$

Hence, a first order approximation results in an error that is quadratic in $\mathrm{u}(\mathrm{r})$. Although it is possible to use higher order Taylor approximations [24], the attempts in this direction have not been conclusive so far.

This work was supported by Huawei.

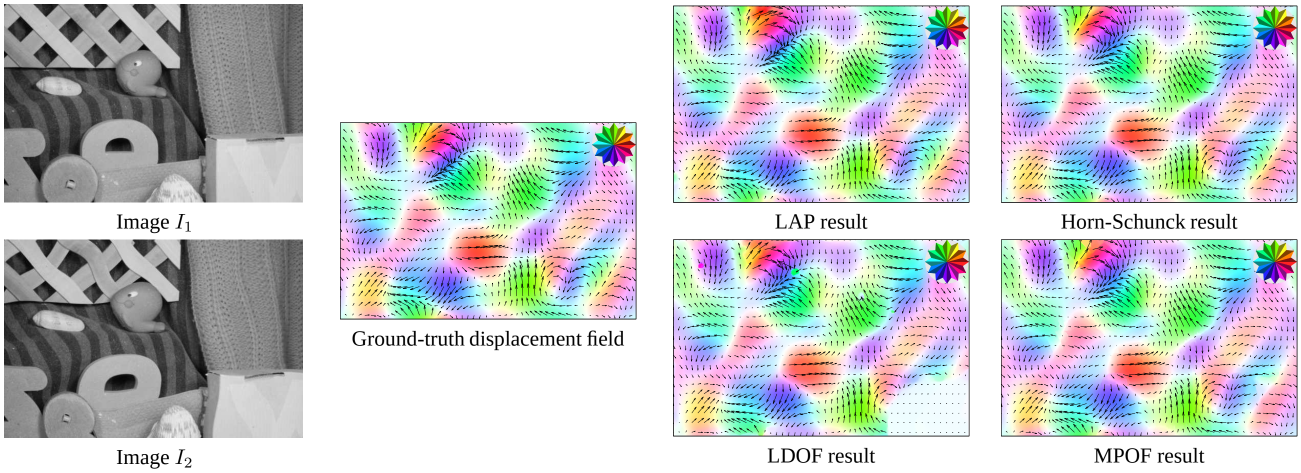

Fig. 1. Synthetic experiment warping image I_1 to image I_2 using a slowly varying displacement field of maximal amplitude 15 pixels. The shown LAP result [19] achieves a median accuracy of 0.010 pixels (mean: 0.102 pixels) in 6 seconds. For comparison, the improved implementation of Horn-Schunck algorithm [18] achieves a median accuracy of 0.604 pixels (mean: 0.868 pixels) in 47 seconds; LDOF [22] achieves a median accuracy of 0.701 pixels (mean: 1.310 pixels) in 30 seconds; and MPOF [23] achieves a median accuracy of 0.623 pixels (mean: 0.964 pixels) in 279 seconds. To facilitate the visual comparisons, we have used a color code to indicate directions (top-right color wheel) and amplitudes, redundantly with the arrows.

Using Fourier variables ($\hat{I}_1(\omega)$ denoting the Fourier transform of $I_1(\mathrm{r})$), $I_1(\mathrm{r}-\mathrm{u}(\mathrm{r}))$ can be expressed as

$$I_1(\mathrm{r}-\mathrm{u}(\mathrm{r})) = \frac{1}{4\pi^2}\int \hat{I}_1(\omega)\mathrm{e}^{-j\mathrm{u}(\mathrm{r})^{\mathrm{T}}\omega}\mathrm{e}^{j\mathrm{r}^{\mathrm{T}}\omega}\,\mathrm{d}\omega \qquad (1)$$

and the Taylor approximation can be seen to derive from the first order Taylor development of the exponential

$$\mathrm{e}^{-j\mathrm{u}(\mathrm{r})^{\mathrm{T}}\omega} = 1 - j\mathrm{u}(\mathrm{r})^{\mathrm{T}}\omega + \mathrm{O}\big(|\mathrm{u}(\mathrm{r})^{\mathrm{T}}\omega|^2\big)$$

A recent approach to optical flow estimation developed by us [19], the local all-pass algorithm, uses a rational approximation (not a polynomial approximation) of the exponential—a Padé approximation. This new algorithm achieves a high accuracy and spatial consistency which makes it outperform the state-of-the-art optical flow algorithms in synthetic experiments. In real-life experiments, the algorithm is still very competitive, although not the best—at least, on some experiments. In addition this algorithm is quite fast (a few seconds for standard 512×512 images).

3. LOCAL ALL-PASS ALGORITHM

The LAP algorithm departs from the observation that, when $\mathrm{u}(\mathrm{r})$ is constant across the image, $I_1(\mathrm{r}-\mathrm{u})$ is exactly the result of the convolution of an all-pass filter, $\delta(\mathrm{r}-\mathrm{u})$, and $I_1(\mathrm{r})$. Hence, the idea is to approximate this ideal filter using an all-pass filter, $h(\mathrm{r})$. It turns out that all-pass filters can always be expressed in the Fourier domain as the ratio

$$\hat{h}(\omega) = \frac{\hat{p}(\omega)}{\hat{p}(-\omega)} \qquad (2)$$

where $p(\mathrm{r})$ is an arbitrary real filter (with a Fourier transform). However, instead of looking for the ideal all-pass filter, the idea developed in the LAP is to approximate the filter $p(\mathrm{r})$ onto a basis of few filters. Then slowly varying flows $\mathrm{u}(\mathrm{r})$ can be estimated by approximating the all-pass filter in local windows. The working principle of the LAP algorithm is that the all-pass filtering relation between the two images can be expressed linearly as a function of $p(\mathrm{r})$:

$$I_2(\mathrm{r}) = h(\mathrm{r}) * I_1(\mathrm{r}) \quad\Leftrightarrow\quad p(-\mathrm{r}) * I_2(\mathrm{r}) = p(\mathrm{r}) * I_1(\mathrm{r}).$$

Then, a simple mean square minimization (fast, non-iterative) provides the parameters representing $p(\mathrm{r})$, from which, a nonlinear accurate formula provides an estimate of the flow $\mathrm{u}(\mathrm{r})$.

Now, the question we want to answer is: if we are able to choose the best all-pass filter $h(\mathrm{r})$ in this constrained framework, what is the order of the approximation of $I_1(\mathrm{r}-\mathrm{u}(\mathrm{r}))$ by $h(\mathrm{r}) * I_1(\mathrm{r})$?

4. PADÉ APPROXIMATION OF THE COMPLEX EXPONENTIAL

To find the approximation order of the LAP algorithm, it is useful to consider Padé approximants of the complex exponential function with equal numerator and denominator degrees [25]. These approximants can be obtained from the continued fraction of e^x [26, p. 70], but we will follow a different approach.

Let us define the sequence of complex functions, $\varepsilon_n(x)$, defined through the recursion

$$\begin{cases} \varepsilon_0(x) = \mathrm{e}^{jx} - 1, \\ \varepsilon_n(x) = j\displaystyle\int_0^x \varepsilon_{n-1}(\xi)(\mathrm{e}^{j(x-\xi)} - 1)\,\mathrm{d}\xi, \ \text{for } n \geq 1. \end{cases} \tag{3}$$

Proposition 1 *The functions $\varepsilon_n(x)$ satisfy the following properties*

i. Sign change: $\varepsilon_n(x) = -\varepsilon_n(-x)\mathrm{e}^{jx}$;

ii. Complex conjugation: $\varepsilon_n(-x)^* = \varepsilon_n(x)$;

iii. Polynomial order: $|\varepsilon_n(x)| \leq 2^{-n}|x|^{2n+1}$.

iv. Taylor: $\varepsilon_n(x) \sim j(-1)^n \frac{x^{2n+1}}{(2n+1)!}$ *as* $x \to 0$

Property iii also implies that $\varepsilon_n(x)$ *is* $\mathrm{O}\big(x^{2n+1}\big)$.

Proof — Properties *i* and *ii*: it is easy to show (using a change of variables $\xi \to -\xi$ in the integral) that $\varepsilon_n(-x)\mathrm{e}^{jx}$ and $\varepsilon_n(-x)^*$ satisfy the same recursion equation as $\varepsilon_n(x)$. Hence, since $\varepsilon_0(-x)\mathrm{e}^{jx} = -\varepsilon_0(x)$ and $\varepsilon_0(-x)^* = \varepsilon_0(x)$, we infer by induction on n that *i* and *ii* are true for all integer $n \geq 0$.

Property *iii*: Thanks to the symmetry *ii*, we can restrict the proof to $x \geq 0$. Using the recursion equation (3), we have the following inequality

$$\begin{aligned} |\varepsilon_n(x)| &\leq \max_{0\leq\xi\leq x} |\varepsilon_{n-1}(\xi)| \int_0^x \underbrace{\left|\mathrm{e}^{j(x-\xi)} - 1\right|}_{\leq x-\xi} \mathrm{d}\xi \\ &\leq \frac{x^2}{2} \max_{0\leq\xi\leq x} |\varepsilon_{n-1}(\xi)| \end{aligned}$$

Since $|\varepsilon_0(x)| \leq x$, we infer that $|\varepsilon_n(x)| \leq 2^{-n}x^{2n+1}$ by induction on n.

Property *iv*: by Taylor, we have $e^{jx} - 1 \sim jx$ as $x \to 0$. The recursion is verified by substitution of $\varepsilon_{n-1}(x) \sim a_{n-1}x^{2n-1}$ into (3) and using the identity

$$a_n x^{2n+1} = -\int_0^x a_{n-1}\xi^{2n-1}(x-\xi)\,\mathrm{d}\xi = \frac{-a_{n-1}x^{2n+1}}{2n(2n+1)}. \ \blacksquare$$

Lemma 1 *There exists a sequence, $P_n(x)$, of real polynomials of degree n such that*

$$\varepsilon_n(x) = P_n(-jx)\mathrm{e}^{jx} - P_n(jx). \tag{4}$$

Proof — We will prove by induction on n that $\varepsilon_n(x)$ can be expressed as $a_n(x)\mathrm{e}^{jx} + b_n(x)$, where $a_n(x)$ and $b_n(x)$ are polynomials of degree n. This property is satisfied for $n = 0$ with $a_0(x) = 1$ and $b_0(x) = -1$. So, let us assume that it is satisfied for some integer $n \geq 0$. We will prove that it will be satisfied for $n + 1$ as well.

By using (3) we find that

$$\begin{aligned} \varepsilon_{n+1}(x) &= j\int_0^x \varepsilon_n(\xi)(\mathrm{e}^{j(x-\xi)} - 1)\,\mathrm{d}\xi \\ &= -\int_0^x E_n(\xi)\mathrm{e}^{j(x-\xi)}\,\mathrm{d}\xi \quad \text{(by parts)} \\ &= -F_n(x)\mathrm{e}^{jx} \end{aligned}$$

where $E_n(x)$ is the primitive of $\varepsilon_n(x)$ that vanishes at 0, and where $F_n(x)$ is the primitive of $E_n(x)\mathrm{e}^{-jx}$ that vanishes at 0.

So, if we assume that $\varepsilon_n(x) = a_n(x)\mathrm{e}^{jx} + b_n(x)$ where $a_n(x)$ and $b_n(x)$ are polynomials of degree n, then its primitive is of the form $E_n(x) = \alpha_n(x)\mathrm{e}^{jx} + \beta_{n+1}(x)$, where $\alpha_n(x)$ is a polynomial of degree n and $\beta_{n+1}(x)$ a polynomial of degree $n + 1$. Then, $F_n(x)$ is the primitive of $\alpha_n(x) + \beta_{n+1}(x)\mathrm{e}^{-jx}$ that vanishes at 0. This function is of the form $a_{n+1}(x) + b_{n+1}(x)\mathrm{e}^{-jx}$ where $a_{n+1}(x)$ and $b_{n+1}(x)$ are polynomials of degree $n + 1$. This shows that ε_{n+1} is of the form $a_{n+1}(x)\mathrm{e}^{jx} + b_{n+1}(x)$.

Then, thanks to the symmetries stated in Proposition 1, we obtain $b_n(x) = -a_n(-x)$ (from property *i*) and that $a_n(x)$ is a real polynomial of the variable jx (from property *ii*); hence, we can choose to define $P_n(-jx) = a_n(x)$. ■

Note: the polynomial sequence $P_n(x)$ can be shown to satisfy the recursion ODE: $-P_n'' + P_n' = P_{n-1}$. It is the only polynomial solution to this equation that satisfies the initial condition $P_n(0) = 2P_n'(0)$ (for $n \geq 1$). For instance, we have

$$\begin{aligned} &P_1(x) = 2 + x, \\ &P_2(x) = 6 + 3x + \frac{x^2}{2}, \\ &P_3(x) = 20 + 10x + 2x^2 + \frac{x^3}{6}, \quad \text{etc.} \end{aligned}$$

As can be observed, the coefficients of these polynomials are strictly positive, a property that can be proven by induction.

Proposition 2 *The polynomials $P_n(x)$ defined in Lemma 1 do not have pure imaginary roots; or, equivalently, if we define $\gamma_n = \inf_{x\in\mathbb{R}} |P_n(jx)|$, then*

$$\gamma_n > 0, \quad \textit{for all positive integer } n.$$

Proof — Let us show that, if for some n, there exists a real x_0 such that $P_n(jx_0) = 0$ then we reach a contradiction. We can assume that $x_0 \neq 0$ because the coefficients of $P_n(x)$ are strictly positive (cf. earlier remark).

First, since $P_n(x)$ is a real polynomial and $x_0 \neq 0$, both jx_0 and $-jx_0$ are roots of $P_n(x)$, which means that $P_n(x)$ can be factorized as $(x^2 + x_0^2)P_{n-2}^1(x)$, where P_{n-2}^1 is a polynomial of degree $n - 2$.

Then, from Proposition 1 (Property *iii* applied to $\varepsilon_n(x)$ of (4)), we know that $P_n(-jx)\mathrm{e}^{jx} - P_n(jx) = O(x^{2n+1})$ which implies that $P_{n-2}^1(-jx)\mathrm{e}^{jx} - P_{n-2}^1(jx) = O(x^{2n+1})$. This is actually impossible, because expressions of the form

$$\varepsilon(x) = P(x)\mathrm{e}^{jx} + Q(x) \tag{5}$$

where $P(x)$ and $Q(x)$ are arbitrary (complex or real) polynomials of degree $m \in \mathbb{N}$ cannot be $O(x^{2m+5})$. To see this, let us perform the following differential operator on the function $\varepsilon(x)$ which we assume to be $O(x^{2m+5})$: $\varepsilon''(x) - j\varepsilon'(x) = \mathrm{e}^{jx}\left\{\mathrm{e}^{-jx}\varepsilon'(x)\right\}'$. Expressing $\varepsilon(x)$ according to (5) we find that

$$\underbrace{\mathrm{e}^{jx}\left\{\mathrm{e}^{-jx}\varepsilon'(x)\right\}'}_{\mathrm{O}(x^{2m+3})} = \underbrace{\big(P''(x)+jP'(x)\big)\,\mathrm{e}^{jx}+Q''(x)-jQ'(x)}_{\text{polynomials of degree } m-1}.$$

The rhs is of the form (5) with m changed into $m-1$ and is now $\mathrm{O}\big(x^{2(m-1)+5}\big)$, so that we can repeat the same differential operator until we obtain polynomials $P(x)$ and $Q(x)$ of degree 0; i.e., constants. Hence, we reach a point where we find that there exist constants p and q such that $p\mathrm{e}^{jx}+q = \mathrm{O}\big(x^5\big)$ which is obviously impossible, since the best order we can get for an expression of the form $p\mathrm{e}^{jx}+q$ is $\mathrm{O}(x)$—reached when $p=-q$. Hence, an expression of the form (5) with polynomials $P(x)$ and $Q(x)$ of degree $m=n-2$ cannot be $\mathrm{O}\big(x^{2n+1}\big)$.

This contradiction shows that our hypothesis on the existence of pure imaginary roots of $P_n(x)$ was wrong. ■

Theorem 1 *A Padé approximation of order $2n$ of the complex exponential function is given by the rational fraction $P_n(jx)/P_n(-jx)$ and we have that*

$$\left|\mathrm{e}^{jx}-\frac{P_n(jx)}{P_n(-jx)}\right| \le \frac{|x|^{2n+1}}{2^n\gamma_n}.$$

This shows that this rational approximation of e^{jx} is $\mathrm{O}\big(x^{2n+1}\big)$.

Proof — We use (4) to get

$$\mathrm{e}^{jx}-\frac{P_n(jx)}{P_n(-jx)} = \frac{\varepsilon_n(x)}{P_n(-jx)}.$$

Then, the theorem results from the inequalities stated in Propositions 1 (Property *iii*) and 2. ■

Note: It is important to notice that, here, the polynomial involved in the rational fraction is only of degree n, despite the fact that the approximation order is twice larger. This is in contrast with polynomial approximations like Taylor's, in which case the order of the approximation is the degree of the approximating polynomial.

5. LAP APPROXIMATION ORDER

We are interested in the order of the approximation of $I_1(\mathrm{r}-\mathrm{u}(\mathrm{r}))$ by $h(\mathrm{r}) * I_1(\mathrm{r})$ when $h(\mathrm{r})$ is an all-pass filter of the form (2). More specifically, like in the LAP algorithm, we assume that the filter $p(\mathrm{r})$ involved in (2) is in the span of a basis of derivatives (up to order n) of a Gaussian function

$$p(\mathrm{r}) = \sum_{l=0}^{n}\sum_{k=0}^{l} a_{k,l}\frac{\partial^l}{\partial x^k \partial y^{l-k}}\left\{\exp\Big(-\frac{x^2+y^2}{2\sigma^2}\Big)\right\}, \quad (6)$$

where σ is a free positive parameter. The cardinality of this basis is $\frac{1}{2}(n+1)(n+2)$, and it is clear that the all-pass filter (2) specified by

$$\hat{p}(\omega) = P_n(-j\mathrm{u}^{\mathrm{T}}\omega)\mathrm{e}^{-\frac{1}{2}\sigma^2\|\omega\|^2}$$

can be expressed on this basis. Typically, in the LAP algorithm, the value chosen for n is either 1 (three basis filters, comprised of up to the first order derivatives), or 2 (six basis filters, comprised of up to the second order derivatives).

Now, we need to introduce a Fourier-based notion of regularity: a function $f(\mathrm{r})$ over $\mathbb{R}^2$ is said to be m times L^1-Fourier differentiable iff both its Fourier transform $\hat{f}(\omega)$ and $\|\omega\|^m\hat{f}(\omega)$ are absolutely integrable. This notion implies—but is not equivalent—that the partial derivatives $\frac{\partial^k f(\mathrm{r})}{\partial x^i \partial y^{k-i}}$ for $0 \le i \le k \le m$ exist and are continuous. Then we have the following theorem.

Theorem 2 *Consider a location r_0 and the local all-pass filter $h_{\mathrm{r}_0}(\mathrm{r})$ defined according to* (2) *with*

$$\hat{p}_{\mathrm{r}_0}(\omega) = P_n(-j\mathrm{u}(\mathrm{r}_0)^{\mathrm{T}}\omega)\mathrm{e}^{-\frac{1}{2}\sigma^2\|\omega\|^2}. \quad (7)$$

Then, if $I_1(\mathrm{r})$ is $(2n+1)$-times L^1-Fourier differentiable (slightly stronger than $\mathrm{C}^{2n+1}(\mathbb{R}^2)$), we have

$$I_1(\mathrm{r}-\mathrm{u}(\mathrm{r}_0)) - h_{\mathrm{r}_0}(\mathrm{r}) * I_1(\mathrm{r}) = \mathrm{O}\big(\|\mathrm{u}(\mathrm{r}_0)\|^{2n+1}\big)\,;$$

i.e., this approximation is of order $2n$.

Proof — We use the inverse Fourier transform formula (1) to get $I_1(\mathrm{r}-\mathrm{u}(\mathrm{r}_0)) - h_{\mathrm{r}_0}(\mathrm{r}) * I_1(\mathrm{r}) =$

$$\frac{1}{4\pi^2}\int \hat{I}_1(\omega)\big(\mathrm{e}^{-j\mathrm{u}(\mathrm{r}_0)^{\mathrm{T}}\omega}-\hat{h}_{\mathrm{r}_0}(\omega)\big)\mathrm{e}^{j\mathrm{r}^{\mathrm{T}}\omega}\,\mathrm{d}\omega.$$

By Theorem 1, we know that

$$\begin{aligned}\left|\mathrm{e}^{-j\mathrm{u}(\mathrm{r}_0)^{\mathrm{T}}\omega}-\hat{h}_{\mathrm{r}_0}(\omega)\right| &\le \text{const}\times|\mathrm{u}(\mathrm{r}_0)^{\mathrm{T}}\omega|^{2n+1}\\ &\le \text{const}\times\|\mathrm{u}(\mathrm{r}_0)\|^{2n+1}\|\omega\|^{2n+1}\end{aligned}$$

where the constant is independent of ω. Hence we can easily bound $\left|I_1(\mathrm{r}-\mathrm{u}(\mathrm{r}_0)) - h_{\mathrm{r}_0}(\mathrm{r}) * I_1(\mathrm{r})\right|$

$$\begin{aligned}&\le \quad \text{const}\times\|\mathrm{u}(\mathrm{r}_0)\|^{2n+1}\int\|\omega\|^{2n+1}|\hat{I}_1(\omega)|\,\mathrm{d}\omega\\ &\le \quad \text{const}'\times\|\mathrm{u}(\mathrm{r}_0)\|^{2n+1}\end{aligned}$$

where the last inequality holds because our L^1-Fourier differentiability assumption on I_1 is equivalent to finiteness of the above integral. ■

6. DISCUSSION

In our current practice [19], LAP is used with $n=1$ (only first order derivatives involved, three basis filters) or $n=2$ (only first and second order derivatives involved, six basis filters). Theorem 2 shows that under a regularity assumption on the image, the LAP algorithm is of approximation order 2 or of order 4. This is remarkable because standard optical flow algorithms are based on a simple first-order approximation of the effect of a displacement—the "optical flow equation". What we have shown in this paper is that, without increasing the differentiation depth, i.e., computing only first order derivatives, and assuming sufficient regularity of the image, we can approximate the effect of a displacement more accurately: the error is a cubic power of the amplitude of the displacement, compared to a quadratic power for the optical flow equation.

7. REFERENCES

[1] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. Black, and R. Szeliski, "A database and evaluation methodology for optical flow," *Int. J. Comput. Vision*, vol. 92, no. 1, pp. 1–31, 2011.

[2] S. Song and R. Leahy, "Computation of 3D velocity fields from 3D cine CT images of a human heart," *IEEE Trans. Med. Imag.*, vol. 10, no. 3, pp. 295–306, 1991.

[3] A Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: A survey," *IEEE Trans. Med. Imag.*, vol. 32, no. 7, pp. 1153–1190, 2013.

[4] J. Barron and A Liptay, "Measuring 3-D plant growth using optical flow," *Bioimaging*, vol. 5, no. 2, pp. 82–86, 1997.

[5] J. Delpiano, J. Jara, J. Scheer, O. Ramírez, J. Ruiz-del Solar, and S. Härtel, "Performance of optical flow techniques for motion analysis of fluorescent point signals in confocal microscopy," *Mach. Vision Applicat.*, vol. 23, no. 4, pp. 675–689, 2012.

[6] B. Horn and B. Schunck, "Determining optical flow," *Artificial Intell.*, vol. 17, no. 1, pp. 185–203, 1981.

[7] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. Int. Joint Conf. Artificial Intell., Vancouver, Canada*, 1981, vol. 2, pp. 674–679.

[8] M. Black and P. Anandan, "The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields," *Comput. Vision Image Understanding*, vol. 63, no. 1, pp. 75 – 104, 1996.

[9] A. Bruhn, J. Weickert, and C. Schnorr, "Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods," *Int. J. Comput. Vision*, vol. 61, no. 3, pp. 211–231, 2005.

[10] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, "High accuracy optical flow estimation based on a theory for warping," in *Proc. European Conf. Comput. Vision (ECCV), Prague, Czech Republic*, 2004, vol. 3024, pp. 25–36.

[11] A. Wedel, T. Pock, C. Zach, H. Bischof, and D. Cremers, "An improved algorithm for TV-L 1 optical flow," in *Stat. Geometrical Approaches to Visual Motion Anal.*, D. Cremers, B. Rosenhahn, A.. Yuille, and F. Schmidt, Eds., vol. 5604 of *Lecture Notes in Computer Science*, pp. 23–45. Springer Berlin Heidelberg, 2009.

[12] W. Dong, G. Shi, X. Hu, and Y. Ma, "Nonlocal sparse and low-rank regularization for optical flow estimation," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4527–4538, 2014.

[13] J.K. Aggarwal and N. Nandhakumar, "On the computation of motion from sequences of images-a review," *Proc. IEEE*, vol. 76, no. 8, pp. 917–935, 1988.

[14] M. Otte and H. H. Nagel, "Optical flow estimation: advances and comparisons," in *Proc. European Conf. Comput Vision (ECCV), Stocklholm, Sweden*, 1994, vol. 1, pp. 51–60.

[15] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of optical flow techniques," *Int. J. Comput. Vision*, vol. 12, no. 1, pp. 43–77, 1994.

[16] A. Mitiche and P. Bouthemy, "Computation and analysis of image motion: a synopsis of current problems and methods," *Int. J. Comput. Vision*, vol. 19, no. 1, pp. 29–55, 1996.

[17] C. Stiller and J. Konrad, "Estimating motion in image sequences," *IEEE Signal Process. Mag.*, vol. 16, no. 4, pp. 70–91, 1999.

[18] D. Sun, S. Roth, and M. Black, "A quantitative analysis of current practices in optical flow estimation and the principles behind them," *Int. J. Comput. Vision*, vol. 106, no. 2, pp. 115–137, 2014.

[19] C. Gilliam and T. Blu, "Local all-pass filters for optical flow estimation," in *Proc. IEEE Int. Conf. Acoust Speech Signal Process. (ICASSP)*, Brisbane Australia, April 2015, to appear.

[20] David J Fleet and Allan D Jepson, "Computation of component image velocity from local phase information," *International journal of computer vision*, vol. 5, no. 1, pp. 77–104, 1990.

[21] David Vernon, "Computation of instantaneous optical flow using the phase of Fourier components," *Image and vision computing*, vol. 17, no. 3, pp. 189–199, 1999.

[22] T. Brox and J. Malik, "Large displacement optical flow: Descriptor matching in variational motion estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 3, pp. 500–513, 2011.

[23] L. Xu, J. Jia, and Y. Matsushita, "Motion detail preserving optical flow estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 9, pp. 1744–1757, 2012.

[24] Hans-Hellmut Nagel, "Displacement vectors derived from second-order intensity variations in image sequences," *Computer Vision, Graphics, and Image Processing*, vol. 21, no. 1, pp. 85–117, 1983.

[25] George A Baker Jr and John L Gammel, "The Padé approximant," *Journal of Mathematical Analysis and Applications*, vol. 2, no. 1, pp. 21–30, 1961.

[26] M Abramowitz and IA Stegun, *Handbook of mathematical functions*, National Bureau of Standards, 1972.