Brightness-compensated 3-D optical flow algorithm for monitoring cochlear motion patterns

Miriam von Tiedemann
Anders Fridberger
Mats Ulfendahl
Karolinska Institutet
Center for Hearing and Communication Research
Department of Clinical Neuroscience
    and
Department of Otolaryngology
SE-171 76 Stockholm, Sweden

Jacques Boutet de Monvel
Institut Pasteur
Unité de Genetique et Physiologie de l'Audition
Inserm UMRS587
Paris, France 75724

Abstract. A method for three-dimensional motion analysis designed for live cell imaging by fluorescence confocal microscopy is described. The approach is based on optical flow computation and takes into account brightness variations in the image scene that are not due to motion, such as photobleaching or fluorescence variations that may reflect changes in cellular physiology. The 3-D optical flow algorithm allowed almost perfect motion estimation on noise-free artificial sequences, and performed with a relative error of <10% on noisy images typical of real experiments. The method was applied to a series of 3-D confocal image stacks from an *in vitro* preparation of the guinea pig cochlea. The complex motions caused by slow pressure changes in the cochlear compartments were quantified. At the surface of the hearing organ, the largest motion component was the transverse one (normal to the surface), but significant radial and longitudinal displacements were also present. The outer hair cell displayed larger radial motion at their basolateral membrane than at their apical surface. These movements reflect mechanical interactions between different cellular structures, which may be important for communicating sound-evoked vibrations to the sensory cells. A better understanding of these interactions is important for testing realistic models of cochlear mechanics. © *2010 Society of Photo-Optical Instrumentation Engineers.* [DOI: 10.1117/1.3494564]

1 Introduction

Sound evokes complex patterns of cellular motion within the hearing organ. A detailed characterization of these motions is essential in order to understand the way cochlear hair cells, the sensory cells of the hearing organ, are stimulated in response to sound.[1,2] The small size, the anatomical location inside thick bone, the fragility of its cells, and the complex three-dimensional (3-D) geometry, combined with a requirement for probing motion at the submicron scale makes it challenging to characterize hearing organ vibrations. Although a number of techniques[3–6] have allowed obtaining a great amount of information, no technique can be said to fully solve the problem.

The most sensitive method is laser interferometry, which allows *in vivo* measurements of cochlear vibrations down to the nanoscale (for reviews, see Refs. 7 and 8). A major limitation of this approach is that it only permits measurements at one point at a time, making it hard to assess cellular deformations and the way different structures are mechanically coupled. An alternative approach is to use optical microscopy combined with motion analysis algorithms. This has permitted a more detailed monitoring of the motion patterns of the hearing organ, but has thus far been restricted to two dimensions.[9–12]

Many algorithms useful for extracting object motion from an image scene have been described (for a review of the older literature, see Ref. 13; more recent developments are reviewed in Ref. 14). A key assumption in many of these algorithms is that brightness changes originate solely from motion. Horn and Schunck[15] formulated this mathematically in the brightness constancy constraint equation [Eq. (1)] which states that the sum of the spatial and temporal derivatives of the image sequence is zero. This implies that all changes in pixel brightness are due to motion alone. Because of detector noise, changes in cellular properties, and in the case of fluorescence experiments, bleaching, this constraint can never be fully satisfied in any biological experiment. Also, in the case of two-dimensional image sequences, the algorithm will only detect the projection of the motion on the image plane. Cells are usually part of 3-D structures that do not respect the two-dimensional (2-D) boundaries of an image. Consequently, it is desirable to develop algorithms that can assess the full 3-D motion pattern while providing a degree of immunity to unavoidable brightness variations. Such an algorithm is described here. The algorithm is an extension of the wavelet-based approach described previously.[16]

Address all correspondence to: Anders Fridberger, M1 Karolinska Universitetssjukhuset Solna, SE-171 76 Stockholm, Sweden; Tel: 46-8-5177-3274; Fax: 46-8-348-546; E-mail: anders.fridberger@ki.se

2 Methods and Materials

2.1 Optical Flow Formulations

Our approach to optical flow estimation is based on applying a differential constancy constraint equation[15] to the images

$$\partial_t I(\mathbf{r},t) + \mathbf{v}(\mathbf{r},t) \cdot \nabla I(\mathbf{r},t) = \partial_t I + v_x \partial_x I + v_y \partial_y I + v_z \partial_z I = 0,$$

(1)

where I denotes image intensity (∂I and ∇I being its temporal derivative and spatial gradient, respectively) and $\mathbf{v} = (v_x, v_y, v_z)$ is the unknown displacement vector, at a particular position $\mathbf{r}$ and time t. For Eq. (1) to be useful, the motion must be sampled at a high enough rate and with a voxel size small enough for the motion vector to vary little among neighboring voxels. However, even under these conditions, Eq. (1) is generally not exactly satisfied in biological experiments. It may however be assumed to be valid after a proper normalization of the image sequence.[17] To implement this idea, let us assume that the image is given by $I(\mathbf{r},t) = m(\mathbf{r},t)\hat{I}(\mathbf{r},t)$, where $\hat{I}(\mathbf{r},t)$ satisfies (1) and $m(\mathbf{r},t)$ is a multiplication factor taking into account intensity variations not caused by motion of objects in the observed scene. The image $I(\mathbf{r},t)$ then satisfies

$$\partial_t I = \partial_t m \hat{I} + m \partial_t \hat{I} = \frac{\dot{m}}{m} I - m \mathbf{v} \nabla \hat{I} = \frac{\dot{m}}{m} I + \frac{\mathbf{v} \cdot \nabla m}{m} I - \mathbf{v} \cdot \nabla I;$$

that is,

$$\partial_t I + v_x \partial_x I + v_y \partial_y I + v_z \partial_z I + \alpha I = 0,$$

(2)

where $\alpha(\mathbf{r},t) = -\partial_t \log m(\mathbf{r},t) - \mathbf{v} \cdot \nabla \log m(\mathbf{r},t)$ is a function of position and time, which we refer to as the brightness variation rate, because it determines the intensity variation of a given image particle along its trajectory. This multiplicative model is convenient for the case of the positive images generated by our confocal microscope, but additive models are, in principle, equivalent and could also be used.

The optic flow components v_x, v_y, v_z and the brightness variation factor α are now to be estimated altogether. To this end, we apply a discrete wavelet transform (DWT) of the image: $I \to \omega_{s,j} = I * \psi_{j,s}$, where $j = 1, \ldots, n_j$ runs over the number of scales used in the DWT, and $s = 0, \ldots, 7$ labels the polarizations of the wavelet filters $\psi_{j,s}$ (there are eight such filters in three dimensions). Equation (2) then yields a set of constraints (one constraint for each wavelet filter),

$$\partial_t \omega_{j,s} + \mathbf{v} \cdot \nabla \omega_{j,s} + \alpha \omega_{j,s} = 0,$$

which hold to a good approximation if the functions $v_x(\mathbf{r},t)$, $v_y(\mathbf{r},t)$, $v_z(\mathbf{r},t)$, and $\alpha(\mathbf{r},t)$ vary little over the supports of the wavelet filters $\psi_{j,s}$. This set of equations can be written in the matrix form

$$A\mathbf{v} + B = 0, \quad \mathbf{v} = (v_x, v_y, v_z, \alpha),$$

(3)

where $A = (A_{js,a})$ and $B = (B_{js})$ are matrices of sizes $n_\omega \times 4$ and $n_\omega \times 1$, respectively (n_ω being the number of wavelet components used), defined by

$$A_{js,a} = \partial_a \omega_{js} \quad a = 1,2,3; \quad A_{js,4} = \omega_{js}; \quad B_{js} = \partial_t \omega_{js}.$$

(4)

The solution to (3) is found by least-squares inversion, i.e., we choose $\mathbf{v}$ to minimize

$$e_{LS}(\mathbf{v}) = \|A\mathbf{v} + B\|^2 = \sum_{j,s} (A_{js,1} v_x + A_{js,2} v_y + A_{js,3} v_z + A_{j,4} \alpha + B_{js})^2.$$

Using matrix notations, the solution (obtained by setting $\partial e_{LS} / \partial \mathbf{v} = 0$) is given by $\mathbf{v} = -(A^T A)^{-1} A^T B$, where A^T and A^{-1} denote the transpose and inverse of the matrix A, respectively. The above least-squares estimation procedure is local in the sense that the total least-squares error $E_{LS}(\mathbf{v}) = \int \|A\mathbf{v} + B\|^2 dx$ is obtained by minimizing the local errors $e_{LS}[v(\mathbf{r},t)] = \|Av(\mathbf{r},t) + B\|^2$ with respect to $\mathbf{v}(\mathbf{r},t)$ independently for each position and time.

In practice, local irregularities sometimes occur even in areas where the aperture problem is expected not to affect results. To reduce such irregularities, the optical flow map was low-pass filtered with a 3-D Gaussian kernel with a standard deviation of $8 \times 8 \times 4$ pixels. For our images, this kernel size turned out to be a good compromise, allowing irregularities to be substantially reduced without spoiling the optical flow accuracy by oversmoothing.

All computations were performed using 64-bit Matlab (the Mathworks, Natick, Massachusetts), running on a Linux PC with 24 GB of random access memory.

2.2 Animal Preparation and Visualization

The preparation used in these experiments has previously been described in detail.[18,11,19] In brief, temporal bones from young anesthetized guinea pigs were excised, using procedures approved by the local ethics committee (permit no. N319/06). The bulla was opened, and the preparation attached to a holder in a chamber containing tissue culture medium (Minimum Essential Medium with Hank's salts, Gibco, Paisley, Scotland). A small opening was made in the basal turn of the cochlea [see schematic drawing in Fig. 1(a)], and a thin piece of plastic tubing inserted in scala tympani (ST) to allow perfusion of oxygenated medium and fluorescent dyes, which were used to label the inner ear structures. The sensory hair cells and auditory nerve fibers were labeled by the styryl dye RH795. Supporting cells were stained by the cytoplasmic dye calcein/acetoxymethyl ester (both from Molecular Probes, Leiden, the Netherlands). An additional opening in the apical part of the cochlea allowed visualization of the organ of Corti while providing exit for the medium.

The preparation was visualized with a water-immersion objective (Achroplan $40\times$, numerical aperture 0.8, Zeiss, Jena, Germany) using a laser-scanning confocal microscope (LSM 510, Zeiss) equipped with a 15-mW krypton/argon laser and a 5-mW helium/neon laser. To avoid photobleaching and consequent cellular damage, all stacks were acquired at the minimum laser power compatible with an acceptable signal-to-noise ratio in the images. In most experiments, this resulted in a laser setting of $\sim 0.25\%$ of the maximum power for the 488-nm laser line and $\sim 10\%$ for the 543-nm line.
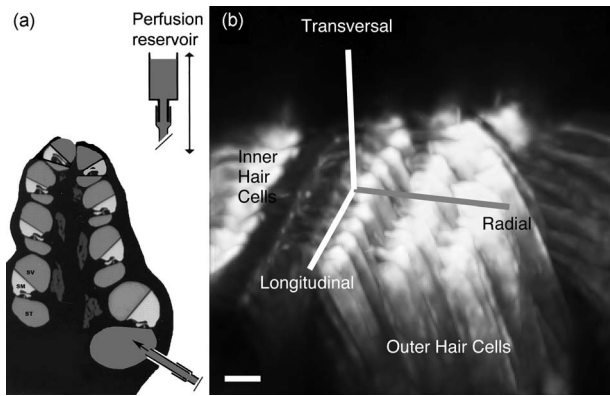
Fig. 1 (a) Drawing showing a cross section of the cochlea and the arrangement for perfusion, with a tube inserted in ST of the basal turn and the reservoir above the fluid level of the chamber. By moving the reservoir, the hydrostatic pressure inside the fluid-filled compartments is shifted; an increased pressure in ST moves the organ of Corti in the direction of *scala vestibuli* (upward in this image). (b) 3-D maximum brightness projection of a stack of images from the apical turn of a guinea pig cochlea. The three orthogonal axes of the standard coordinate system are superimposed on the image. The longitudinal axis is aligned with the first row of outer hair cells, pointing toward the base of the cochlea. The radial axis points from the center of the cochlea in the direction of the outer hair cells, and the transversal vector is direct to the radial-longitudinal plane. Scale bar, 10 μm.

2.3 Organ of Corti Motion Evoked by Pressure Changes

ST pressure was changed by moving the perfusion reservoir above and below the fluid level of the preparation in a sequence forming a cycle, typically 0, +10, 0, −10, 0 cm, where 0 cm corresponded to the surface of the liquid surrounding the preparation. Each position was maintained for a few seconds to allow the hearing organ to reach a new equilibrium position, and a stack of confocal images acquired. The typical pixel format of the stacks was $512 \times 512 \times 32$, requiring about 40 s to acquire at a pixel dwell time of 6.4 μs. The spacing between the sections was 0.5–1 μm.

The pressure changes in the ST are linearly related to the changes in height of the reservoir. They remain within the physiological range and may be assumed a fraction of 1 Pa or less.[20] These changes induce reproducible, micron-sized movements of the cochlear partition.

In these experiments, the orientation of the cochlea varied and to compare results across experiments, it is necessary to use a standardized coordinate system. The reference frame that we used is an orthogonal coordinate system whose three axes at any particular point along the cochlea are defined as follows. The *x*-axis, or longitudinal axis, is parallel to the coil of the cochlea and points toward its base; the *y*-axis, or radial axis, is taken parallel to the plane of the reticular lamina and oriented from the inner hair cells to the outer hair cells; the *z*-axis, or transverse axis, is taken orthogonal to *x* and *y*. Figure 1(b) shows the coordinate system superimposed on a maximum brightness projection of a stack of images acquired at different focal planes within the organ of Corti.

3 Results

3.1 Validation of the Optical Flow Algorithm

The optical flow algorithm was evaluated on two different synthetic image sequences, one showing a sine pattern and the other a more complex template image showing cochlear structures. Both images had a maximal amplitude of 1 and minimum amplitude of zero. A uniform translation was applied to each pattern, using bicubic interpolation. The exact motion is therefore known, and a direct measure of the error in the optical flow estimation is possible. Because noise is unavoidably present in real image sequences, Gaussian or Poisson-distributed noise was added to each frame following interpolation. No systematic difference between the two types of noise was noted. The Gaussian noise had a mean of zero and standard deviation (σ) ranging between 0.01 and 0.04.

Results are illustrated in Figs. 2(a)–2(c) for the sine sequence. The standard deviation of the Gaussian noise in Fig. 2(a) was 2.5%, representing an image with noise levels typical of those found experimentally. In the absence of added noise, the magnitude error, which is defined as the difference in Euclidean vector magnitude of the estimated and true motion vectors, was found to be <10% for all displacements of <8 voxels [Fig. 2(b) solid line]. A systematic downward translation of the entire curve was observed when the noise level increased. At intermediate noise levels, this downward translation resulted in curves with <10% magnitude error throughout the range tested.

It is also important to determine the direction of motion. The angular error [Fig. 2(c)] is a measure of the difference in orientation of the estimated and true motion vectors in the plane that they span. The angular error was found to be <10 deg at all tested noise levels and displacements. For this artificial image, the direction of motion was therefore more precisely estimated than its size.

The above computations were performed on an image sequence that lacked variations in brightness over time. To simulate an actual experiment with bleaching due to laser exposure, additional tests were performed on sequences where the image amplitude was reduced by 2% for each motion step. After interpolation and brightness reduction, Gaussian random noise was added as described above. The results are shown in Figs. 2(d)–2(f). The left half of Fig. 2(d) is identical to the one shown in Fig. 2(a). To give a measure of the magnitude of the brightness variations, the right half of Fig. 2(d) shows the final image in the series, where brightness was reduced to 56% of the original. Apparently, the algorithm can handle such brightness variation. The main effect is a reduction in the slope of the magnitude error curves [Fig. 2(e)], which results in larger magnitude errors at high noise levels. However, angular errors were almost unchanged [Fig. 2(f)]. All the curves remain <10 deg, except for large displacements at the highest noise level. To estimate the improvement from using the brightness-compensated algorithm, we also performed computations where brightness compensation was removed from the algorithm. When using sequences with the above reduction in brightness, we found a substantially reduced performance for small motions (<3 voxels): The magnitude errors increased and so did angular errors. However, for larger displacements, the results were similar to those obtained with the brightness-compensated algorithm. The per-
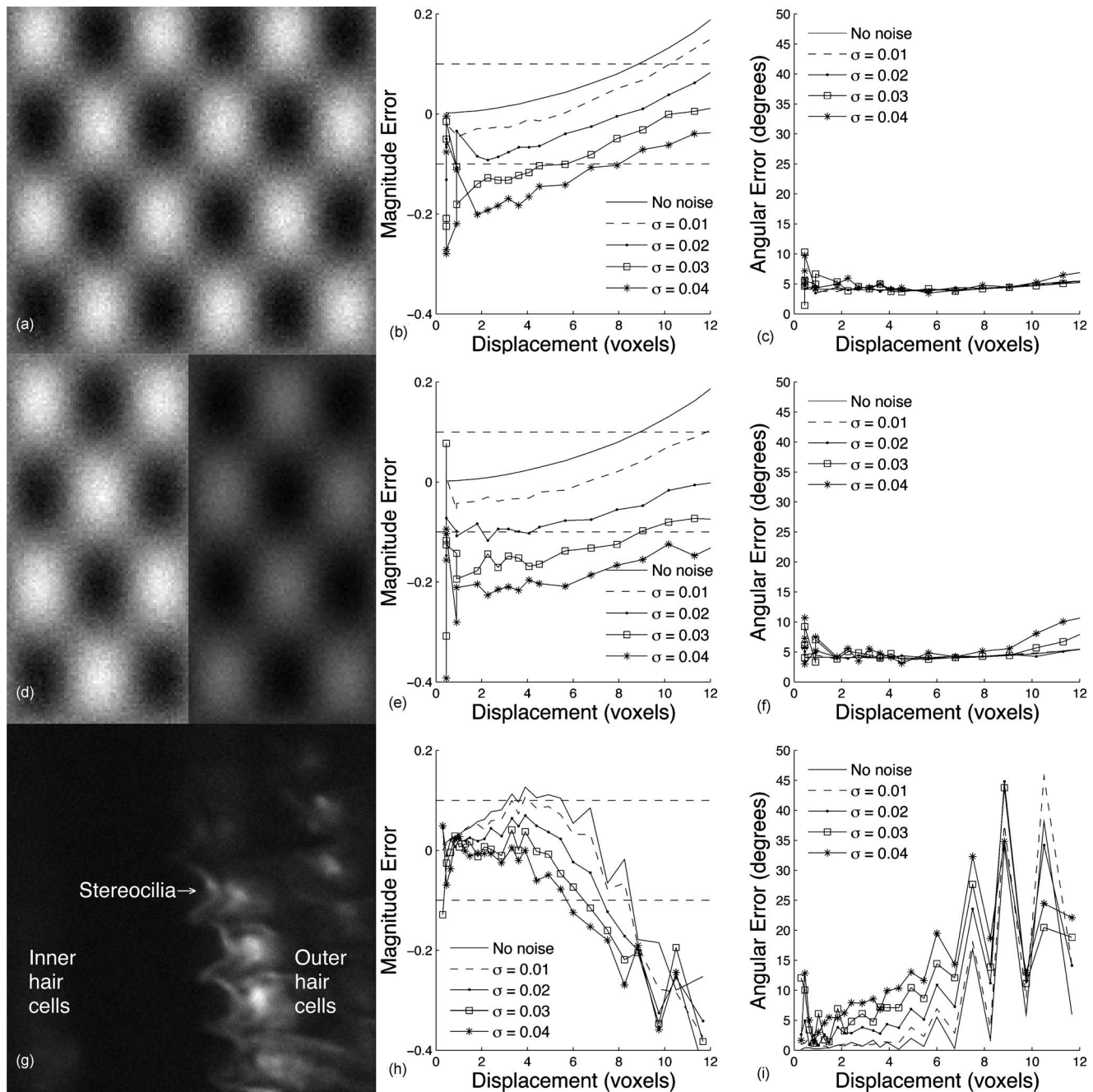
Fig. 2 Motion tracking on two different sequences of 3-D test images. Motion was generated through bicubic interpolations (a)–(c). An artificial sine pattern was used. No brightness variation is present, but noise added as indicated by the legend inside panels (b) and (c) (d)–(f). Identical image sequence as in (a)–(c) but the overall brightness was reduced by 2% for each motion step. Following brightness reduction, noise was added as in (a)–(c) (g)–(i). An image stack from the hearing organ was subjected to the same manipulations as the sine image in panel (d)–(f). The magnitude error is defined as the root-mean-square distance between the observed and the real motion vectors, relative to the size of the real motion. The angular error is obtained by taking the arccos of the dot product of the real and estimated 3-D motion vectors.

formance did not depend on the pattern of brightness reductions: A linear decrease in brightness over time gave the same results as a monoexponential one.

The moving sine pattern is an unrealistic test case because it contains only a single spatial frequency. Sharp boundaries are absent, making this sequence difficult for any gradient-based method. A more realistic test is achieved by using real images, which contain larger spatial variations in brightness.

Figure 2(g) shows a group of hair cells *in situ*. Important cellular structures such as the stereocilia bundles at the top of the outer hair cells are clearly visible. Noise levels, brightness variations and imposed motions were identical to those used in Fig. 2(d)–2(f). From Fig. 2(h), it is evident that small displacements are detected with high fidelity when noise levels are low ($\sigma < 0.02$), but there is an overestimation of motion
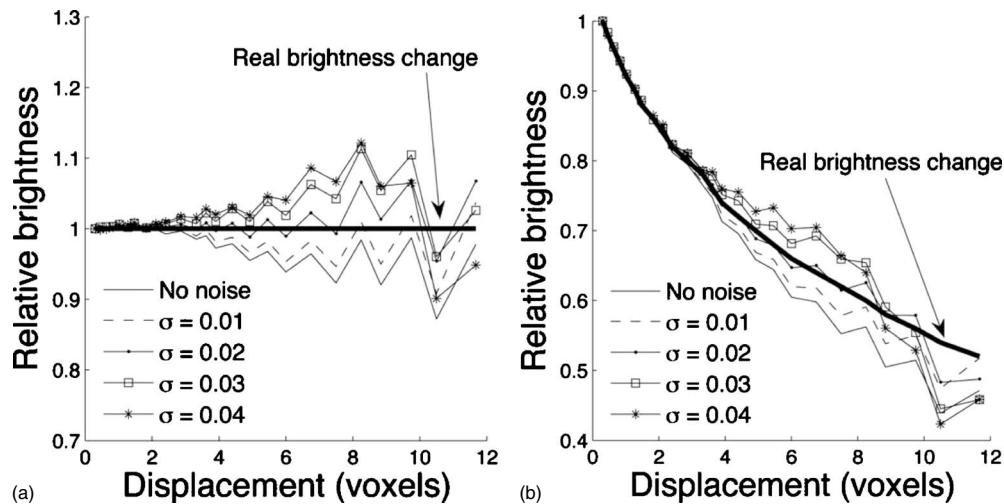
Fig. 3 (a) Pixel brightness estimation in the absence of a true change in brightness. The outer hair cell sequence was used. The real change in brightness is given by the thick line. Noise levels were identical to Fig. 2. (b) Same computation as in (a), but this time, a 2% reduction in brightness was present for each motion step.

for the range between 3 and 5 voxels. At higher noise levels, good motion detection is seen at <5 voxels, but results deteriorate for displacements larger than this. Thus, in terms of the magnitude error, some noise seems to produce better results than no noise. The reason for this unusual, but quite desirable behavior is not currently known. However, the angular error behaves differently. Near-perfect estimation of vector directions was achieved for artificial motions that lacked noise. Results get systematically worse with increasing noise, but the error remains at <10 deg at all noise levels if the motion is <4 voxels.

The factor α in Eq. (2) provides an estimate of the brightness changes in the image sequence. To evaluate the accuracy of this estimate, we first assessed its variability in image sequences that lacked variations in brightness. In this case, the algorithm returned a faithful estimate, provided that motion was <5 voxels [Fig. 3(a)]. The most precise estimates were obtained when a moderate amount of noise was present. The "no-noise" condition as well as the extreme case of noise with a standard deviation of 0.05 both caused performance deterioration. The brightness estimates therefore appear to behave similarly to the motion estimates in Fig. 2(h). When actual brightness variations were present, a similar pattern was seen [Fig. 3(b)]. The estimated brightness tracked the real brightness change, and the best estimates were obtained at intermediate noise levels.

3.2 Validation of the Experimental Setup

To assess the stability of our experimental setup, we estimated the random motion of the table and scanning mirrors of the confocal microscope by imaging a test sample repeatedly under identical conditions. The sample was a slide containing roots from *convallaria majalis*, which have a stable structure with bright fluorescence. A series of 11 three-dimensional image stacks were acquired, at two different magnifications, with a 60-s pause between each acquisition. Note that averaging was not used, and that the pixel dwell time was kept brief, resulting in images with substantial noise [Fig. 4(a)].

The values for the stage motion given below therefore represent worst-case estimates because the noise in the images will lead to increased errors when calculating motion.

The root-mean-square displacement per step, θ_1, estimated from np selected trajectories on the image, was computed as

$$\theta_1 = \sqrt{\frac{1}{np}\sum_{i=1}^{np}\left(\frac{1}{n}\sum_{j=1}^{n}x_{i,j}^2\right)},$$

where n is the number of steps and $x_{i,j}$ represents the j'th displacement along a given trajectory i. Similarly, the mean cumulative displacement after n steps is defined by
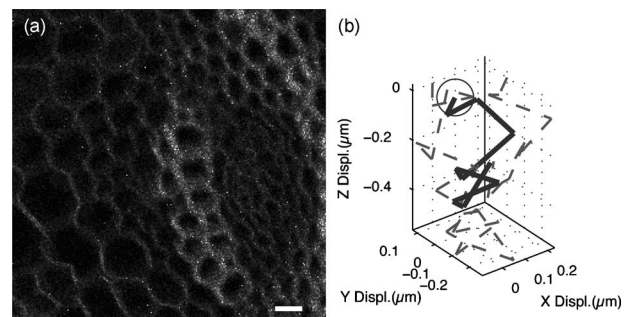

Fig. 4 (a) Sample image showing the fluorescent roots of *convallaria majalis*. Note the relatively high noise level of the image. Scale bar, 20 μm. (b) Random motion of the sample shown in (a) on the microscope stage. The graph shows the 3-D optical flow trajectory of the stage (black solid trace) averaged over 10 image points; the broken lines are projected trajectories on *xy*, *yz*, and *xz* planes. Note the random character of this trajectory, reminiscent of Brownian motion. The ring marks the starting point of the motion trajectory.

Table 1 Estimates of the step size along the trajectory shown in Fig. 4. n_p is the number of image points over which averages are taken; θ_1 is the root-mean-square displacement of the stage for a duration of 60 s; θ_n is the cumulative rms displacement after a duration of $n \times 60$ s, here for $n=10$; $n_{eff}=(\theta_n/\theta_1)^2$ is to be interpreted as an effective number of steps, which should be approximately equal to n if the stage was undergoing unconstrained Brownian motion. The small n_{eff} indicates that the stage is physically constrained and does not wander around its mean position by more than a fraction of a micron.

np	θ_1(m)	θ_n(m)	n_{eff}
10	1.8575×10^{-07}	4.5189×10^{-07}	5.9185
40	2.3148×10^{-07}	4.9177×10^{-07}	4.5134
60	2.2013×10^{-07}	4.7400×10^{-07}	4.6368

$$\theta_n = \sqrt{\frac{1}{np}\sum_{i=1}^{np} R_i^2},$$

where R_i is the cumulative displacement after n steps over the trajectory i.

The displacement of the scanned volume was estimated with the optical flow algorithm. As seen in Fig. 4(b), the motion of the microscope stage from one stack to the next appeared to be random, with a root-mean-square displacement per step $\theta_1=0.19$ to 0.23 μm and a cumulative displacement after 10 steps of $\theta_{10} \approx 0.5$ μm (pixel size 0.9 μm). Note that motion along the Z-axis, the position of which changes during acquisition of a focus series, was most pronounced. This probably relates to slight imperfections in the focus motor controlling the position of the microscope's table. We may define a quantity, $n_{eff}=(\theta_n/\theta_1)^2$, which is interpreted as an effective number of steps, which should be approximately equal to n if the stage was undergoing unconstrained Brownian motion. In Table 1, it is seen that n_{eff} is substantially smaller than n, which indicates that the stage and scanning system is in fact physically constrained. Obviously, this is a very desirable feature. During 10 min, the system does not wander around its mean position by more than a fraction of 1 μm, a prerequisite for its use for motion measurements. Note that this constitutes an upper bound on stage motion, as this value will be influenced by any error present in the optical flow calculation in addition to the actual motion of the stage. We conclude that errors caused by random shifts within the microscope were small and affected the optical flow estimates made in the cochlear preparation by less than a few percent.

3.3 Optical Flow Estimate of 3-D Organ of Corti Motion

The main motivation for the development of our optical flow algorithm was to analyze the 3-D motion patterns of the complex cellular structures of the hearing organ in response to different kinds of mechanical stimulation, including sound. As a first step, we analyzed the 3-D displacement evoked by quasi-static ST pressure changes. Such pressure changes were used in several previous studies;[20,18] this type of stimulation leads to a consistent motion pattern within the organ of Corti, the magnitude of which depends on the quality of the seal between the cochlea and the perfusion tubing, the size of the opening at the apex, and the ST impedance.

Seven different experiments showing pressure change–induced movements of outer hair cells in the apical part of the cochlea were analyzed using identical parameters in the optical flow algorithm. An example showing data from two preparations is given in Fig. 5. As seen in Figs. 5(a) and 5(b), the orientation of these two preparations with respect to the optical axis of the microscope differed substantially. The optical section in Fig. 5(a) is well aligned with the long axis of the cylindrical outer hair cells, whereas the preparation in Fig. 5(b) was nearly perpendicular to this, with the short axis of the hair cells nearly parallel to the image plane. The optical flow vectors were therefore converted to the standard coordinate system shown in Fig. 1(b) where the transverse axis is perpendicular to the reticular lamina, the radial axis points
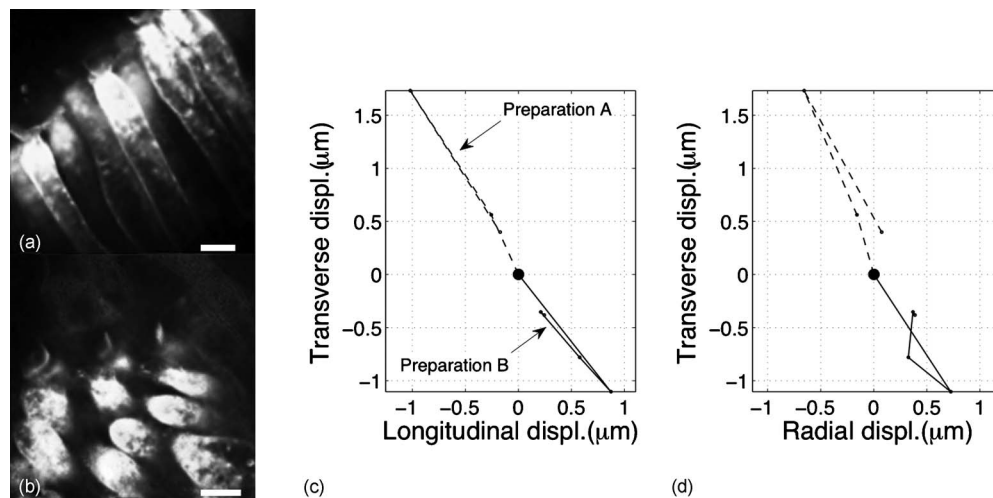

Fig. 5 Pressure-evoked motions of the organ of Corti *in situ*. (a) and (b) show outer hair cells from two different preparations, the longitudinal and radial motions of which are displayed in (c) and (d), respectively. The preparation shown in (a) was subjected to positive scala tympani pressure changes while negative pressures were applied to the preparation shown in (b). Scale bars, 10 μm.
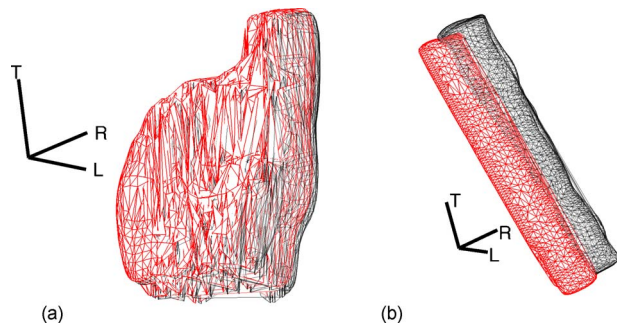
Fig. 6 3-D reconstructions of (a) an inner hair cell and (b) an outer hair cell. The original positions of the cells are given by the red wire frames, which is then deformed according to the computed optical flow map, to generate the gray wire frame that represents the final position of the cell. Scale bars denote a 10-μm distance along each of the three principal axes, transversal (T), radial (R), and longitudinal (L). (Color online only.)

away from the center of the cochlea, and the longitudinal axis is directed along the cochlear spiral, toward its base.

Positive ST pressures were used in Fig. 5(b). The resulting motions of the apex of a first row outer hair cell are plotted with dashed lines in Fig. 5(c) and 5(d). The transverse displacement, given on the Y-axis, was largest, with amplitude of 1.5 μm. There was also a smaller motion component directed to the apex of the cochlea, along the cochlear spiral. This longitudinal component reached an amplitude of 1 μm. As seen in Fig. 5(d), the smallest motion was directed to the center of the cochlea. In the preparation shown in Fig. 5(b), negative ST pressures were used. Despite the very different orientation of this preparation, the main axis of the motion vectors is quite similar to the ones described above, and the transverse component remains the largest one. This gives additional confidence in the optical flow algorithm. Evidently, the motion of the organ of Corti evoked by these pressure changes is complex, with significant components in all three directions.

Confocal microscopy and optical flow makes it possible to assess the deformation of cells inside the hearing organ. Image data were used to generate the 3-D wire-frame models shown in Fig. 6. Figure 6(a) shows an inner hair cell with its characteristic pear-shaped cell body, and its apical surface perpendicular to the transversal axis. The original wire-frame model of the cell (red) was deformed according to the computed optical flow map. This generated a second wire frame, shown in gray. Following a pressure change, motion components in all three directions were seen, but the largest component was in the radial direction, and the cell therefore appears to be translated toward the right of the image. Figure 6(b) shows data from an outer hair cell in the same preparation as Fig. 6(a). The two cells were close to each other, but not at exactly the same longitudinal position. Note the different scales used in Figs. 6(a) and 6(b). At the outer hair cell, a small motion component in the transversal direction was present, but the largest motion was radial. This radial motion is much larger than that found at the inner hair cells. The radial motion was larger near the bottom of the cell, which therefore shows a swinging motion, but little apparent deformation. By comparing the data in Figs. 5 and 6, it is apparent

that motions at the surface of the hearing organ, displayed in Fig. 5, are different from those found inside (Fig. 6). In particular, the cell bodies of outer hair cells make sideways motions that are larger than those seen at the surface. These sideways motions were bigger than the surface motion in six out of seven cases.

4 Discussion

In this paper, we present a new method for brightness-adjusted 3-D optical flow calculation based on the discrete wavelet transform. Brightness adjustment is highly desirable because bleaching is a common experimental problem in confocal microscopy. Cells may also undergo physiologically relevant changes that alter their fluorescence. Such changes could severely limit the capacity for motion detection, unless a brightness-compensated algorithm is used.

Our algorithm was validated with artificial as well as real image sequences that were degraded by noise. These tests show that the performance of the algorithm depends on the properties of the image, a factor that needs to be considered when performing experiments. Obviously, an image with poor contrast, high noise levels, and small gradients will result in suboptimal motion estimation. Efforts should therefore be made to control noise and, because errors increase for motions of >6 voxels, to tailor the voxel size to the expected magnitude of the motion. Fortunately, both of these factors can to a large extent be controlled on current confocal microscopes.

A fundamental physical problem that cannot be circumvented is that brightness changes originating from motion cannot be fully separated from other types of brightness changes (e.g., Ref. 21). In practice, this creates a trade-off between motion detection capabilities and brightness change detection. The primary purpose of our algorithm is motion detection, and as seen in the performance evaluations in Fig. 2, the algorithm performs very well even under conditions of elevated noise.

4.1 Hearing Organ Motion

Optical flow algorithms have previously been used in several studies in the context of hearing research.[9,10,16,22,23] Their use arises naturally from the fact that the hearing organ is a complex structure, the function of which depends on interactions between different cell types. Probing such interactions requires a system capable of at least 2D measurements.

The quasi-static pressure changes investigated in the current study represent a first step toward investigating sound-evoked cellular interactions in the hearing organ. We show here that the motion evoked by such pressure changes is complex, with components in all three directions. At the surface of the hearing organ, the reticular lamina, the transversal motion component was the largest one, followed by the longitudinal one. The longitudinal component was directed toward the apex of the cochlea. This may be a consequence of the anatomical arrangement of cells within the tissue: The Deiters cells are supporting cells with long processes extending in the apical direction from the base of the outer hair cells. This results in longitudinal mechanical coupling that may promote motion directed at the apex, in the forward direction of the traveling wave.

Interestingly, the cell bodies of outer hair cells showed larger motion than those seen at the surface. A similar behavior was observed in a previous study using 2-D optical flow, where motion was provoked by pressure gradients much larger than those utilized here.[18] This behavior also occurs during sound stimulation,[10] and during electrical stimulation of isolated cochleae, where movements inside the hearing organ were substantially larger than those seen at the surface.[24] Collectively, these results suggest that radial movements of outer hair cell bodies contribute substantially to the internal mechanics of the organ of Corti.

Acknowledgments

References

1. J. Ashmore, "Cochlear outer hair cell motility," *Physiol. Rev.* **88**, 173–210 (2008).
2. R. Fettiplace and C. M. Hackney, "The sensory and motor roles of auditory hair cells," *Nat. Rev. Neurosci.* **7**, 19–29 (2006).
3. G. von Bekesy, *Experiments in Hearing*, McGraw-Hill, New York (1960).
4. B. M. Johnstone and A. J. Boyle, "Basilar membrane vibration examined with the Mössbauer technique," *Science* **158**, 389–390 (1967).
5. S. M. Khanna and D. G. B. Leonard, "Basilar membrane tuning in the cat cochlea," *Science* **215**, 305–306 (1982).
6. N. Choudhury, G. Song, F. Chen, S. Matthews, T. Tschinkel, J. Zheng, S. L. Jacques, and A. L. Nuttall, "Low coherence interferometry of the cochlear partition," *Hear. Res.* **220**, 1–9 (2006).
7. L. Robles and M. A. Ruggero, "Mechanics of the mammalian cochlea," *Physiol. Rev.* **81**, 1305–1352 (2001).
8. M. Ulfendahl, "Mechanical responses of the mammalian cochlea," *Prog. Neurobiol.* **53**, 331–380 (1997).
9. X. Hu, B. N. Evans, and P. Dallos, "Direct visualization of organ of Corti kinematics in a hemicochlea," *J. Neurophysiol.* **82**, 2798–2807 (1999).
10. A. Fridberger and J. Boutet de Monvel, "Sound-induced differential motion within the hearing organ," *Nat. Neurosci.* **6**, 446–448 (2003).
11. A. Fridberger, I. Tomo, and J. Boutet de Monvel, "Imaging hair cell transduction at the speed of sound: Dynamic behavior of mammalian stereocilia," *Proc. Natl. Acad. Sci. U.S.A.* **103**, 1918–1923 (2006).
12. I. Tomo, J. Boutet de Monvel, and A. Fridberger, "Sound-evoked radial strain in the hearing organ," *Biophys. J.* **93**, 3279–3284 (2007).
13. J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of optical-flow techniques," *Int. J. Comput. Vis.* **12**, 43–77 (1994).
14. D. J. Fleet and Y. Weiss, "Optical flow estimation," in *Handbook of Mathematical Models in Computer Vision*, N. Paragios, Y. Chen, and O. Faugeras, Editors, pp. 239–258, Springer, New York (2006).
15. B. K. P. Horn and B. G. Schunck, "Determining optical-flow," *Artif. Intell.* **17**, 185–203 (1981).
16. A. Fridberger, J. Widengren, and J. Boutet de Monvel, "Measuring hearing organ vibration patterns with confocal microscopy and optical flow," *Biophys. J.* **86**, 535–543 (2004).
17. S. Negahdaripour, "Revised definition of optical flow: Integration of radiometric and geometric cues for dynamic scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.* **20**, 961–979 (1998).
18. A. Fridberger, J. Boutet de Monvel, and M. Ulfendahl, "Internal shearing within the hearing organ evoked by basilar membrane motion," *J. Neurosci.* **22**, 9850–9857 (2002).
19. M. Ulfendahl, S. M. Khanna, A. Fridberger, Å. Flock, B. Flock, and W. Jäger, "Mechanical response characteristics of the hearing organ in the low-frequency regions of the cochlea," *J. Neurophysiol.* **76**, 3850–3862 (1996).
20. A. Fridberger, J. T. van Maarseveen, E. Scarfone, M. Ulfendahl, B. Flock, and Å. Flock, "Pressure-induced basilar membrane position shifts and the stimulus-evoked potentials in the low-frequency region of the guinea pig cochlea," *Acta Physiol. Scand.* **161**, 239–252 (1997).
21. H. W. Haussecker and D. J. Fleet, "Computing optical flow with physical models of brightness variation," *IEEE Trans. Pattern Anal. Mach. Intell.* **23**, 661–673 (2001).
22. A. J. Aranyosi and D. M. Freeman, "Sound-induced motions of individual cochlear hair bundles," *Biophys. J.* **87**, 3536–3546 (2004).
23. H. Cai, C. P. Richter, and R. S. Chadwick, "Motion analysis in the hemicochlea," *Biophys. J.* **85**, 1929–1937 (2003).
24. K. D. Karavitaki and D. C. Mountain, "Imaging electrically evoked micromechanical motion within the organ of Corti of the excised gerbil cochlea," *Biophys. J.* **92**, 3294–316 (2007).