

Review

Not peer-reviewed version

---

# Optical Flow-Based Algorithms for Real-Time Awareness of Hazardous Events

---

[Stiliyan Kalitzin](#)\*, [Simeon Karpuzov](#), [George Petkov](#)

Posted Date: 8 October 2025

doi: 10.20944/preprints202510.0431.v1

Keywords: optical flow; tracking; convulsive seizures; apnea; falls



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

# Optical Flow-Based Algorithms for Real-Time Awareness of Hazardous Events

Stiliyan Kalitzin <sup>1,2,\*</sup>, Simeon Karpuzov <sup>3</sup> and George Petkov <sup>3</sup>

<sup>1</sup> Stichting Epilepsie Instellingen Nederland (SEIN), Achterweg 5, 2103 SW Heemstede, The Netherlands

<sup>2</sup> Image Sciences Institute, University Medical Center Utrecht, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands

<sup>3</sup> GATE Institute, Sofia University, 1164 Sofia, Bulgaria

\* Correspondence: skalitzin@sein.nl Tel.: +31 623693619

## Abstract

Safety and security are major priorities in modern society. Especially for vulnerable groups of individuals, such as the elderly and patients with disabilities, providing a safe environment and adequate alerting for debilitating events and situations can be critical. Wearable devices can be effective but require frequent maintenance and can be obstructive or stigmatizing. Video monitoring by trained operators solves those issues but requires human resources, time and attention and may present certain privacy issues. We propose optical flow-based automated approaches for a multitude of situation awareness and event alerting challenges. The core of our method is an algorithm providing the reconstruction of global movement parameters from video sequences. This way the computationally most intensive task is performed once and the output is dispatched to a variety of modules dedicated to detect adverse events such as convulsive seizures, falls, apnea and signs of possible post-seizure arrests. The software modules can operate separately or in parallel as required. Our results show that the optical flow-based detectors provide robust performance and are suitable for real-time alerting systems. In addition, the optical flow reconstruction is applicable to real-time tracking and stabilizing video sequences. The proposed system is already functional and undergoes field trials for cases of epileptic patients.

**Keywords:** optical flow; tracking; convulsive seizures; apnea; falls

---

## 1. Introduction

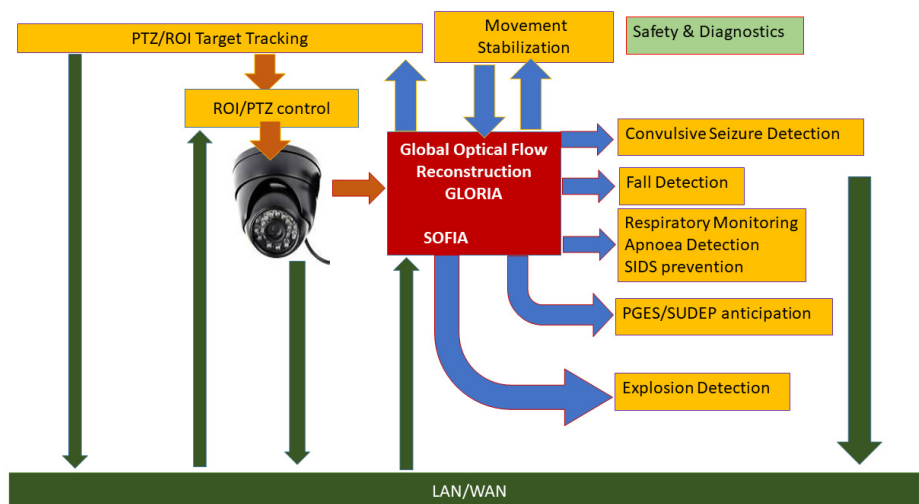
The main objective of this review is to present a common algorithmic approach to a variety of real-time video observation challenges. These challenges arise from clinical and general practice scenarios where video observation can be critical for the safety and security of the monitored population. We will further address each of the scenarios and events separately and here we first introduce the generic idea of optical flow (OF) image and video processing.

OF is a powerful technique [1–3] that gives the reconstruction of object displacements from analyzing related pairs of images where the objects are recorded. Although the most common use is in inferring the velocities of moving objects from video sequences, it has also been successfully applied in stereo vision [4,5] for reconstructing depth information from the disparity between the images provided by two (or more) spatially separated cameras. There are numerous algorithms available but most of them use only the intensity information in the images and ignore the spectral content, the color. We have designed our own proprietary algorithm in [6] where all spectral components of the image sequences participate in parallel in the reconstruction process. In addition, our method, named SOFIA from Spectral Optical Flow Iterative Algorithm, provides iterative multi-scale reconstruction of the displacement field. The spatial scale or aperture parameter has been studied comprehensively earlier [3–5,7]. Our approach however, goes one step further and uses iteratively a sequence of scales, running from coarse-grained to fine, in order to stabilize the solution

of the inverse problem without losing spatial resolution. Such multi-scale method provides hierarchical control over the level of detail that is needed for each individual application, ranging from large-scale global displacements to finer, pixel-level ones.

For global displacements, where individual pixels are not relevant, the reconstruction of the OF and subsequent aggregation of the velocity field is obviously a highly redundant procedure. For such applications we have developed a second proprietary algorithm [8], named GLORIA, where global displacements, such as translations, rotations, dilatations, shear or any other group transformations can be reconstructed directly without solving the OF problem at the pixel level. Such an approach assumes certain knowledge, or model behind the OF content, but it has significant computational advantages that allow usage in real-time applications. It gives as an output the group parameter variations that explain the differences in the sequences of images.

Figure 1 illustrates the overall spectrum of the OF applications reviewed here. For the majority of tasks, we refer to GLORIA global reconstruction. The latter is best suitable in scenarios where the overall behavior is relevant for detection or alerting and the exact localization of the process is not required. These are the cases for monitoring convulsive epileptic seizures, falls, respiratory disruptions, object tracking and image stabilization. For the application of detection and localization of explosions, we use the SOFIA algorithm. Here we briefly introduce the individual implementation modules and challenges.



**Figure 1.** The generic scheme of using optical flow reconstruction results in various application modules. Camera streaming input (USB or IP connections) is used for the estimation of the global movement rates (GLORIA algorithm) or the local velocity vector field (SOFIA algorithm). The global parameters can be sent in parallel to an array of modules each providing specific alerts, or tracking and stabilizing functionalities as indicated in the orange boxes. Only for the purposes of explosion detection, localization and charge estimation, the SOFIA algorithm is enrolled. Tracking can be realized either by dynamic region of interest (ROI) or PTZ camera control as provided by the hardware (USB or IP interface). Blue arrows indicate exchange of data between software modules, brown arrows represent direct hardware connections, such as USB and green arrows symbolize generic TCP/IP connectivity used for larger-scale server/cloud-based implementations.

The principal motivation for developing our OF remote detection techniques was the need for remote alerting of **major convulsive seizures** in patients with epileptic condition. Epilepsy is a debilitating disease of the central nervous system [9] that can negatively affect the lives of those suffering from it. There are various forms and suspected causes [10] for the condition, but in general, epilepsy manifests with intermittent abnormal states, fits or seizures that interrupt normal behavior. Perhaps the most disrupting types of epileptic seizures are the convulsive ones where the patient falls into uncontrollable oscillatory body movements [11,12]. During these states the individual is

particularly vulnerable and at higher risk of injuries or even death. Especially hazardous are terminal cases of Sudden Unexpected Death in Epilepsy, or SUDEP [13,14]. The timely detection of epileptic seizures can therefore be essential for protecting the life of the patients in certain situations [15]. Because of the sudden, unpredictable occurrence of the epileptic seizures, a continuous monitoring of the patients is essential for their safety. Automated detection of seizures has long been studied [16–19] and effective techniques based on electroencephalography (EEG) signals are now in use in specialized diagnostic facilities. Those systems are however not directly applicable for home or residential facilities use as they require trained technicians to attach and control the EEG electrodes. The latter can also cause discomfort to the patient. Wearable devices that use 3D accelerometers are available and validated for use in patients [20–24]. Although effective and reliable, these devices need constant care, charging and proper attachment. They may, therefore, not be the optimal solution for some groups of patients. Their visible presence may also pose ethical issues related to stigmatization. Alternatively, bed mounted pressure or movement detectors are also used [25,26] but their effectiveness can be hampered by the position of the patient and the direction of the convulsive movements. Notably, both classes of the above-mentioned detectors rely on limited measures of movements from one single spatial point. These shortcomings can be resolved by using video observation that can provide a “holistic” view of the whole of substantial part of the patient’s body. Continuous monitoring by operators however, is a time and attention consuming process demanding great amounts of operator work [20–24] load. In addition, privacy concerns may restrict or even prevent the use of manned video monitoring. To address these issues, automated video detection techniques have been investigated [27–36]. In these works, recorded video data has been used to analyze the movements of the patient and validate the detection algorithms. Such systems can be useful as tools for offline video screening and will increase the efficiency of the clinical diagnostic workflow. It is not always clear however, which of the proposed algorithms are suitable for real-time alerting of convulsive seizures.

In our work [37] we reported results from operational system for real time continuous monitoring and alerting. It employs the GLORIA OF reconstruction algorithm and is in use in a residential care facility. In addition, the system allows for continuous, on-the-fly, personalization and adaptation of the algorithm parameters [38] by using an unsupervised learning paradigm. With this functionality, the alerting device finds optimal balance between specificity and sensitivity and can adjust its operational modalities in cases of changes of the environment or patient’s status.

In addition to detecting of convulsive epileptic seizures, we investigated the possibility of predicting **post-ictal generalized electrographic suppression** events (PGES) that may be a factor in the SUDEP cases [39]. In [40] we found, using spectral and image analysis of the OF, that in cases of tonic-clonic convulsive motor events, the frequency of the convulsions or the body movements per second, exponentially decreases towards the end of the seizure. We also developed and validated an algorithm for automated estimation of the rate of the decrease from the video data. Based on a hypothesis derived from a computational model [41], we related the amount of decrease of the convulsive frequency to the occurrence and the duration of a PGES events. This finding was further validated on cases with clinical PGES [42] and may provide a method for diagnosing and even alerting in real-time of possible post-ictal arrests of brain activity.

Another area of application of real-time optical flow video analysis is the detection and alerting for **falls**. Falls are perhaps the most common causes of injuries, especially among the elderly population [43–46]. Also, in the vulnerable population of epileptic patients, falls resulting from epileptic arrests can be a major complication factor [47–49]. Accordingly, a lot of research and development has been dedicated to the detection and prevention of these events [50–56]. The challenge of robust detection of falls has led to accumulating of empirical data in natural and simulated environments [57,58] and the development of new algorithms [59–65]. One of the major challenges is the reliable distinction of fall events from other situations in real-world data [66] and the comparison of the results to simulated scenarios [67]. As with the alerting for epileptic seizures, wearable devices provide solution [61,68] but also have their functional and support limitations. Non-

wearable fall detection systems [69] have been also developed and implemented, including approaches based on sound signals [70–72] produced by a falling person.

Possibly the most reliable and studied fall detection systems are based on automated video monitoring [73–82]. Algorithms based on depth imaging [83], some using Microsoft Kinect stereo vision device, are also proposed [84]. Notably there are few works addressing the issue by combining multiple modalities [85]. The simultaneous use of video and audio signals has been found to improve the performance of the detector [86,87]. Recently, machine learning paradigms have been added to the detection techniques offering personalization of the methods [88–93]. Optical Flow is one of the widely spread methods for detecting falls in video sequences [89,94,95]. We applied our proprietary global motion reconstruction algorithm GLORIA in [93] where the six principal movement components are fed into a pre-trained convolutional neural network for classification. Such an approach allows including a fall alerting module to our integral awareness concept.

One of the potential causes of death during or immediately after epileptic seizures is **respiratory arrest, or apnea** [96]. Together with cardiac arrests [14] this may be a major confounding factor in the cases of SUDEP. While in cases of epileptic condition seizure detection can be the lead safety modality [97], the detection and management of apnea events for the general population is relevant as well [98,99]. The cessation of breathing is the most common symptom for the Sudden Infant Death Syndrome (SIDS) that usually occurs during sleep and the cause often relates to breathing problems.

Devices dedicated to apnea detection during sleep have been proposed and tested in various conditions. Especially relevant are methods based on non-obstructive contactless sensor modalities [100–103] including sensors inbuilt in smart phones [104]. A depth registration method using Microsoft Kinect sensor has also been investigated [105]. Perhaps the most challenging approaches for apnea detection and alerting are those using life video observations. Cameras are now available in all price ranges and they are suitable for day and night continuous monitoring of subjects. To automate the task of recognizing apnea events from video images in real time, researchers have developed effective algorithms. Numerous approaches have been proposed [106–113] in the literature. A common feature in these works is the tracking of the respiratory chest movements of the subject [114]. In our work [115], we applied global motion reconstruction of the video optical flow and subsequent time-frequency analysis followed by classification algorithms to identify possible events of respiratory movement arrests. In a recent patent application [US20230270337A1], tracking of respiratory frequency provides effective method for alerting of SIDS.

Optical flow reconstruction at pixel scale [6] was also used in the context of detection and quantification of explosions in public spaces [116]. Fast cameras registering images in time-loops provided view from multiple locations. Dedicated algorithm for 3D scene reconstruction was constructed to localize point events registered simultaneously from the individual cameras. This part of the technique goes outside the scope of the present work. The optical flow analysis, together with the reconstructed depth information, provided an estimate of the charge of the explosion. Explosion events were detected and quantified from the local dilatation component calculated as the divergence of the velocity vector field at suitable spatial scale. Further in the methods we give some more details of this concept, here we note that optical flow-based velocimetry has also been explored for near-field explosion tracking [117].

The last two topics of this survey concerns indirect application of the optical flow global motion reconstruction. The first application is dedicated to **automated tracking** of moving objects or subjects [118–121]. This is achieved by either defining a dynamic region of interest (ROI) containing the object or by applying physical camera movements such as pan, tilt and zoom (PTZ). This is valuable addition to the monitoring paradigms described above, as manual object tracking is an extremely labor intensive and attention demanding process. Automated tracking in video sequences has been extensively investigated especially for applications related to traffic management and self-driving vehicles [122–126] or surveillance systems [127,128]. Methods dedicated to human movements in behavioral tasks has also been reported [129] in applications where the objectives are mainly related to the challenge of computer-human interfaces [130–132]. In our approach published in [133] and in

a filed patent application, we used the global movement parameters reconstruction GLORIA to infer the transformation of a ROI containing the tracked object. Leaving the technical description for the next paragraph, we note that OF based methods have been introduced in other works [134,135], however no use of the direct transformation parameter reconstruction has been made. To compare, our approach reduces the computational load and makes possible the implementation of the algorithm in real-time. In addition to the single camera-tracking problem, simultaneous monitoring from several cameras has been in the focus of interest of researchers [136–142]. We have addressed the multi-camera tracking challenge by adding adaptive algorithms [143] that reinforce the interaction between the individual sensors in the course of the observation process. Deep learning paradigm has been also employed [144] in multi-camera application for traffic monitoring. In our approach, the coupling between the individual camera tracking routines is constantly adjusted according to the correlations between the OF measurements. We have studied both linear correlation couplings and non-linear association measures. In this way, we have established a dynamic paradigm for video OF based sensor fusion reinforcement. The fusion between multiple sensor observations is a general concept that can be employed in a broader array of applications [145–147].

Finally, we introduce the application of GLORIA method to **stabilizing of video sequences** [US 2022/0207657] when artefacts from camera motion are present. Although optical flow techniques have been used earlier for stabilizing camera imaging [148–150], our approach brings two essential novel features. First, it uses the global motion parameters, namely translation, rotation, dilatation and shear directly reconstructed from the image sequence and therefore avoid the computationally demanding pixel-level reconstruction of the optical flow. Next, we use the group properties of the global transformations and integrate the frame-by-frame changes into an aggregated image transformation. For this purpose, the group laws of vector diffeomorphisms are applied as we explain later in the methods.

The rest of the paper is organized as follows. In the Materials and Methods section, we give the basic formulations of the methods used for the different tasks graphically presented as blocks in Figure 1. We start with the definition of our proprietary SOFIA and GLORIA optical flow algorithms. Next, the application of the GLORIA output for detecting convulsive seizures, falls and apnea adverse events is explained. The extension of the seizure detection algorithm to post-ictal suppression forecast is also presented. Explosion detection, localization and charge estimation from optical flow features is briefly explained. At the end of the methodological section, we focus on the use of global motion optical flow reconstruction for tracking objects and for stabilizing video sequences affected by camera movements.

In the Results section, we report our major findings from our work on the variety of applications. Some limitations and possible extensions of the methodology presented in the Discussion section. An overall assessment of the use of the proposed approaches is offered in the Conclusions.

## 2. Materials and Methods

In the next two subsections, we present in short, the methods introduced in our works [6,8].

### 2.1. Spectral Optical Flow Iterative Algorithm (SOFIA)

Here we recall the well-known concept of optical flow. A deformation of an object or a media can be described by the change of its positions according to some deformation parameter  $t$  (time in the case of temporal processes):

$$x(t + \delta t) = x(t) + v(x, t)\delta t \quad (1)$$

In (1)  $v(x, t)$  is the vector field generating the deformation with an infinitesimal parameter change  $\delta t$  which in the case of motion sequences is the time incremental step. We denote a multi-channel image registering a scene as  $L^c(x, t); c = 1, \dots, N_c, x \in R^2$  (to simplify the notations we consider the channels to be a discrete set of spectral components or colors). If no other changes present

in the scene, the image will change according to the “back-transformation” rule, i.e., the new image values at given point are those transported from the old one due to the spatial deformation.

$$L^c(x, t + \delta t) = L^c(x - v(x, t)\delta t, t) \quad (2)$$

Optical flow reconstruction is then an algorithm that attempts to determine the deformation field  $v(x, t)$  given the image evolution. Assuming small changes and continuous differentiable functions we can rewrite Equation (2) as a differential equation:

$$\begin{aligned} \frac{dL^c}{dt} &= -\nabla L^c \cdot v \equiv \nabla_v L^c; \\ \nabla_v &\equiv v \cdot \nabla \equiv \sum_k v_k \nabla_k; \nabla_k L^c \equiv \frac{\partial L^c}{\partial x_k} \end{aligned} \quad (3)$$

We use here notations from differential geometry where the vector field is a differential operator  $\nabla_v$ . From Equation (3) it is clear that in the monochromatic case  $N_c = 1$  the deformation field is defined only along the image gradient and the reconstruction problem is underdetermined. On the contrary, if  $N_c > 2$  the problem may be over-determined as the number of equations will exceed the number of unknown fields (here and throughout this work we assume two spatial dimensions only although generalization to higher image dimensions is straightforward). However, if the spectrum is degenerate, for example when all spectral components are linearly dependent, the problem is still under-determined. To account for both under and over-determined situations we first postulate the following minimization problem defined by the quadratic local cost-function in each point  $(x, t)$  of the image sequence as:

$$\begin{aligned} C\{L^c(x, t), v(x, t)\} &\equiv \sum_c \left| \frac{dL^c(x, t)}{dt} + \nabla L^c(x, t) \cdot v(x, t) \right|^2 \\ v(x, t) &= \operatorname{argmin}_v [C\{v(x, t)\}] \end{aligned} \quad (4)$$

Clearly because the cost-function in Equation (4) is positive and the solution for  $v(x, t)$  always exists. However, this solution may not be unique because of possible zero modes, i.e., local directions of the deformation field along which the cost-functional is invariant.

Applying the stationarity condition for the minimization problem (1) and introducing the quantities:

$$H_k = -\sum_c \frac{dL^c}{dt} \nabla_k L^c; S_{kj} = \sum_c \nabla_k L^c \nabla_j L^c; j, k = 1, 2 \quad (5)$$

The equation for the velocity vector field minimizing the function is:

$$\sum_j S_{kj}(x, t) v_j(x, t) = H_k(x, t); \quad (6)$$

In definition (2)  $S_{kj}$  will be referred as the structural tensor and  $H_k$  as the driving vector field.

In some applications, it might be advantageous to look for smooth solutions for the optical flow equation. To formulate the problem, we modify the cost function so that in each Gaussian neighborhood of the point  $x$  on the image, the optical flow velocity field is assumed to be the spatially constant vector that can “explain” best the averaged changes in the image evolution in this neighborhood. Therefore, we can modify (literally blur or smoothen) the quadratic cost function (3) in each point  $x$  of the image and its neighborhood as:

$$C^\sigma\{v(x)\} \equiv \sum_y G(x, y, \sigma) \sum_c \left| \frac{dL^c(y, t)}{dt} + \nabla L^c(y, t) \cdot v(x, t) \right|^2 \quad (7)$$

Where the Gaussian kernel is defined as:

$$G(x, y, \sigma) = \frac{1}{N_\sigma} e^{-\frac{(x-y)^2}{\sigma^2}} \quad (8)$$

In Equation (8) the normalization factor  $N_\sigma$  is conveniently chosen to provide unit area under the aperture function. Applying the stationarity condition to the so postulated smoothed cost function leads to the modified equation:

$$S_{kj}^\sigma(x, t) v_j(x, t) = H_k^\sigma(x, t) \quad (9)$$

The smoothed structural tensor and driving vector are obtained as:

$$H_k^\sigma(x, t) \equiv \sum_y G(x, y, \sigma) H_k(y, t); \quad (10)$$

$$S_{kj}^\sigma(x, t) \equiv \sum_y G(x, y, \sigma) S_{kj}(y, t)$$

We can now invert the Equation (9) to obtain explicit unique solution (we skip here the introduction of a regularization parameter, leaving this to the original work) for the optical flow vector field, for a given scale:

$$v_j(x, t) = (S_{kj}^\sigma(x, t))^{-1} H_k^\sigma(x, t) \quad (11)$$

Let denote the solution as a functional of the image and its deformation, the scale parameter as:

$$v_j(x, t) = v_j\{L^c(x, t + \delta t), L^c(x, t), \sigma\} \quad (12)$$

We can approach now the task of finding a detailed optical flow solution by iteratively solving the optical flow equation for a series of  $\sigma^n < \sigma^{n-1} < \dots < \sigma^1$  decreasing scales using the solution of each coarser scale to deform the image and use it as input for obtaining the optical flow at the next finer scale. The iterative procedure can be expressed by the following iteration algorithm:

$$v^{(1)}(x, t) = v\{L^c(x, t + \delta t), L^c(x, t), \sigma^1, \rho\}$$

$$v^{(k+1)}(x, t) = v^{(k)}(x - w^{(k)}(x, t)\delta t, t) + w^{(k)}(x, t); \quad (13)$$

$$k = 1 \dots (n - 1)$$

$$w^{(k)}(x, t) \equiv v\{L^c(x, t + \delta t), L^c(x - v^{(k)}(x, t)\delta t, t), \sigma^{k+1}, \rho\}$$

The last iteration produces an optical flow vector field  $v^{(n)}(x, t)$  representing the result of zooming-down through all scales.

## 2.2. Global Lie-Algebra Optical Flow Reconstruction Algorithm (GLORIA)

In some applications, it might be advantageous to look first or only for solutions for the optical flow equation that represent known group transformations.

$$v(x, t) \equiv \sum_a A^u(t) v^u(x), u = 1 \cdots N_G \quad (14)$$

In Equation (5)  $v^u(x)$  are the vector fields corresponding to each group generator and  $A^u(t)$  are the corresponding transformation parameters, or group velocities in the case of velocity reconstruction context. We can then reformulate the minimization problem by substituting (5) into the cost-function (4) and consider it as a minimization problem for determining the group-coefficients  $A^u(t)$ .

$$C\{A\} \equiv \sum_{c,x} \left| \frac{dL^c(x,t)}{dt} + \sum_u A^u(t) v^u(x) \cdot \nabla L^c(x,t) \right|^2 \quad (15)$$

$$A(t) = \operatorname{argmin}_A [C\{A(t)\}]$$

Using notations from differential geometry, we can introduce the generators of infinitesimal transformations algebra as a set of differential operators.

$$G^u(x) \equiv \sum_k v_k^u(x) \nabla_k \quad (16)$$

The operators defined in (16) form the Lie algebra of the transformation group.

Applying the stationarity condition for the minimization problem (4) and introducing the quantities:

$$H^u = - \sum_{x,k,c} v_k^u(x) \frac{dL^c}{dt} \nabla_k L^u; \quad (17)$$

$$S^{uq} = \sum_{x,k,j,c} v_k^u(x) \nabla_k L^c \nabla_j L^c v_j^q(x); u, q = 1 \cdots N_A$$

The equation for the coefficients minimizing the function is:

$$\sum_q S^{uq} A^q = H^u \quad (18)$$

We can now invert Equation (18) to obtain the unique solution (we skip again the regularization step in case of singular matrix  $S^{uq}$ ) for the optical flow vector field coefficients defined in (14):

$$A^q = \sum_u (S^{uq})^{-1} H^u \quad (19)$$

We apply the above reconstruction method in sequences of two-dimensional images, restricting the transformations to the six parameters non-homogeneous linear group:

$$\begin{aligned} G^{\text{translations}_1}(x) &= \nabla_1; \quad G^{\text{translations}_2}(x) = \nabla_2; \\ G^{\text{rotation}}(x) &= x_2 \nabla_1 - x_1 \nabla_2; \quad G^{\text{dilatation}}(x) = x_1 \nabla_1 + x_2 \nabla_2; \\ G^{\text{shear}_1}(x) &= x_1 \nabla_2; \quad G^{\text{shear}_2}(x) = x_2 \nabla_1 \end{aligned} \quad (20)$$

Those are the two translations, rotation, dilatation and two shear transformations.

### 2.3. Detection of Convulsive Epileptic Seizures

Applying the algorithm GLORIA on the image sequence with the generators (20) produces six time series representing the rates of changes (group velocities) of the six two-dimensional linear inhomogeneous transformations.

$$L^c(x, y, t); c = \{R, G, B\} \Rightarrow V_g(t); g = \{TrX, TrY, Rot, Dil, ShX, ShY\} \quad (21)$$

We use next a set of Gabor wavelets (normalized to unit 1-norm and zero mean) with exponentially increasing wavelengths  $f_k, k = 1..200$

$$W_g(t, f_k) = \left| \int_{t'} dt' g(t - t', f_k) V_g(t') \right|$$

$$g(t - t', f) = (e^{-\pi^2 \alpha^2 f^2 (t-t')^2 - i2\pi f(t-t')} - O_f) / N_f$$

$$W(t, f_k) \equiv \langle W_g(t, f_k) \rangle_g; \quad (22)$$

$$W^q(f_k) \equiv \langle W(t, f_k) \rangle_{t \in q}$$

$$f_k = f_{min} e^{(k-1)\mu}$$

For the exact definitions and normalizations, we refer to earlier publications [26,35] and here we note that the wavelet spectrum in (22) is a time-average along each images sequence window denoted with  $q$ .

Next, we define the “epileptic content” as the fraction of the wavelet energy contained in the frequency range defined here as  $[f_a, f_b]$ .

$$E(q, f_a, f_b) \equiv \frac{\sum_{f \in [f_a, f_b]} W^q(f)}{\sum_f W^q(f)} \quad (23)$$

In the “rigid” application, as well as an initial setting for the adaptive scheme, we use the default range of  $f \in [f_a, f_b] \equiv [2 \dots 7]$  Hz that represents the most common observed frequencies in convulsive motor seizures. To compensate for different frequency ranges that may be used, we calculated also the same quantity in (3) but for a signal with “flat” spectrum representing random noisy input. Then we rescale the epileptic marker as

$$\check{E}(q, f_a, f_b) = \frac{E(q, f_a, f_b) - E_0(f_a, f_b)}{1 - E_0(f_a, f_b)} \quad (24)$$

Here  $E_0$  is the relative wavelet spectral power of a white noise. Note that in (23) and (24) the quantity  $q$  is a discrete index representing the frame sequence number and corresponds, as stated earlier, to a time window conveniently chosen of approximately 1.5 seconds.

We use three parameters  $[N, n, T]$  to detect an event (seizure alert) in real time. At each time instance, we take the seizure marker (3a) in the  $N$  preceding windows. If from those  $N$ , at least  $n$  have values  $\check{E} > T$  an event is generated and eventually (if within the time selected for alerts) sent as an alert to the observation post. The default values are  $[7 \ 6 \ 0.4]$ . This corresponds to a criterion that detects if from the past 10.5 seconds at least 9 seconds contain epileptic “charge” (24) higher than 0.4. These values are used in the rigid mode as well as an initial setting in the adaptive mode described in [37,38].

The design and operation of the adaptive algorithm, representing a reinforcement learning approach, goes beyond the scope of this work. We notice only that the proposed scheme adjusts both the frequency range and the detection parameters  $[N, n, T]$  while performing the detection task. A clustering algorithm applied on the optical flow global movement traces (21) provides the labeling

used for a “ground truth”. We refer to [38] for a detailed description of the unsupervised reinforcement learning technique.

#### 2.4. Forecasting Postictal Generalized Electrographis Suppression (PGES)

The change in clonic frequency during a convulsive seizure can be modelled by fitting a linear equation to the logarithm of the inter-clonic interval [41]. If the times of successive clonic discharges for a given seizure are  $t_k$  (marked for example by visual inspection of the EEG traces or in video recordings), then exponential slowing down can be formulated as:

$$ICI_k \equiv t_{k+1} - t_k = C_0 e^{\alpha \tau_k}; \tau_k \equiv (t_{k+1} + t_k)/2 \quad (25)$$

In Equation (25),  $\alpha$  is a constant defining the exponential slowing. Our hypothesis, validated in [41] is that the overall effect of slowing-down is a factor correlated to the PGES occurrence and duration. The total effect of ictal slowing for each seizure is quantified as

$$ICI_{term} \equiv C_0 e^{\alpha T_{seizure}} \quad (26)$$

In the above definition  $ICI_{term}$  is the terminal inter-clonic interval, the  $C_0$  and  $\alpha$  parameters are derived for each case from the linear fit procedure in Equation (25), and  $T_{seizure}$  is the total duration of the seizure.

The optical flow technique was used to estimate the parameters of the ictal frequency decrease [40]. Starting point is the Gabor spectrum  $W(t, f_k)$  as defined in Equation (22). Because of the exponential increase of the wavelet central frequencies, an exponential decrease of the clonic frequency will appear in a straight line in the time-spectral image. We use this fact to estimate the position and slope of such a line by applying two-dimensional integral Radon transformation, performing integration along all lines in the time-frequency space:

$$RW(r, \theta) \equiv \int_{-inf}^{+inf} W(t(u), f(u)) du \quad (27)$$

$$[t(u), f(u)] \equiv [(u \sin \theta + r \cos \theta), (r \sin \theta - u \cos \theta)]$$

We applied further a simple global maximum detection for the  $RW(r, \theta)$  function and determine the angle and distance parameters of the dominant ridge line as

$$\{r_M, \theta_M\} = \text{armax}\{RW(r, \theta)\} \quad (28)$$

Finally, one can estimate from (22) and (25) the exponential constant  $\alpha$  as:

$$\alpha = \mu \tan^{-1}(\theta_M) \quad (29)$$

The above estimate was done for multiple video recordings of convulsive seizures and used to establish associations with the PGES occurrence and duration. For more technical and analytic insight, we refer to [40].

#### 2.5. Detection of Falls

Here we present briefly only the essential parts of the fall detection algorithm originally introduced in [88]. In this original work, standard optical flow pixel-level algorithm was used but the general methodology, apart from some spatial correction factors, is applicable to the new GLORIA technique.

Assuming a position of the camera that laterally registers the space of observation, the motion component from the set (20), (21) relevant for detection of falls is the vertical velocity corresponding to the translational component  $V_y(t)$  as function of time. Taking a discrete derivative of this time series, we can calculate the vertical acceleration  $A_y(t) = V_y(t) - V_y(t-1)$ . We can assume also that positive values of  $V_y$  correspond to downward motion (otherwise we can invert the sign of the parameter).

From the functions  $A_y(t)$  and  $V_y(t)$  we define a triplet of time-series features  $\{A(t_A), V(t_V), D(t_D)\}$  corresponding to the local positive maxima of the functions  $\{A_y(t), V_y(t), -A_y(t)\}$ . These features are the maximal downward acceleration, velocity and deceleration. An event eligible for fall detection is characterized by these three features whenever the positions of the consecutive maxima are ordered as  $t_A < t_V < t_D$ .

In addition to the three optical flow derived features  $\{A, V, D\}$  we use a fourth one associated with the sound that a falling person may cause. For each window (we use windows of 3 seconds with two seconds overlap, step of one second) we calculate the Gaussian smoothed Hilbert envelope of the audio signal (aperture of 0.1 second) and take the ratio of the maximal to the minimal values. The ratio  $S$  is then associated to all events in the corresponding window where the detection takes place. This way we obtain four features  $\{A, V, D, S\}$  to classify an event as a fall.

The rest of the algorithm involves training technique, support vector machine (SVM) with a radial basis function kernel, to establish the domain in the four-dimensional feature space corresponding to fall events.

In reference [93] we propose a variant of this algorithm that employs all six reconstructed global movements and is enhanced with alternative machine learning techniques such as convolutional neural networks (CNN), skipping however the audio signal component. This approach is less sensitive to the position of the camera and avoids synchronization and reverberation problems associated with the audio registration in real world settings.

## 2.6. Detection of Respiratory Arrests, Apnea

Following the methodology of [115], the same optical flow reconstruction and Gabor wavelet spectral decomposition is used as with the convulsive seizure detection, repeating steps (21), (22). The relative spectrum essential for respiratory tracking is defined similarly to (23) but without the window-averaging of the spectra:

$$RE(t, f_a, f_b) \equiv \frac{\sum_{f \in [f_a, f_b]} W(t, f)}{\sum_f W(t, f)} \quad (30)$$

Where now  $[f_a, f_b] = [0.1, 1] \text{ Hz}$ . The denominator in Equation (30) is the total spectrum for all wavelet central frequencies (0.08 to 5 Hz in this implementation):

$$TS(t) = \sum_f W(t, f) \quad (31)$$

We note that up to this point the algorithm can be modularly linked to the seizure detection processing using the same computational resources.

The specific respiratory arrest events are detected by further post-processing of relative and total spectra defined in (30) and (31).

First, we define a range of 200 scales  $s_k$  ( $k = 1, 2, \dots, 200$ ), with exponentially spaced values in the range 25-500 pixels. For each scale, an aperture sigmoid template is defined for window  $\tau$ :

$$S(\tau, k) = N_k^{S-1} \frac{e^{\tau/s_k} - e^{-\tau/s_k}}{e^{\tau/s_k} + e^{-\tau/s_k}} e^{-|\tau|^2/s_k^2}; \tau = [-3s_k : 3s_k] \quad (32)$$

together with the Gaussian aperture template:

$$G(\tau, k) = N_k^{G^{-1}} e^{-|\tau|^2/s_k^2} \quad (33)$$

In Equations (32) and (33) L2 normalization was applied through the coefficients  $N_k^{S,G^{-1}}$ , with  $N_k$  defined as the squared sum of the  $k$ th aperture template. The time window in (32) and (33) is chosen to be of three scale lengths, as values outside this range are suppressed by the Gaussian aperture factor. Sigmoid time-scale modulation  $m$  can then be obtained using the convolutions between the filters and the RE signal:

$$m(t, k) = \frac{\int_{-\infty}^{\infty} S(t - \tau) RE(\tau) d\tau}{\int_{-\infty}^{\infty} G(t - \tau) RE(\tau) d\tau} \quad (34)$$

To quantify the presence of significant respiratory range power drops, we calculated the mean sigmoid modulation  $M$  over the scales that correspond to observed drop times:

$$M(t) = \langle m(t, k) \rangle_{k \in s_{drop}} \quad (35)$$

Drop times were observed to be between 4.0 and 8.2 seconds in test recordings, and correspond to filters  $s_{drop} \in [s_{70}, s_{129}]$ .

Potential respiratory events are defined at the times of local positive maximums of  $M(t)$ ,  $t_M = \{t: M(t) > \max(M(t-1), M(t+1)); M(t) > 0\}$ . The first feature to be used for apnea detection, the sigmoid modulation maximum is then:

$$SMM(t_M) = M(t_M) \quad (36)$$

A second classification feature quantifying the change of total power at the time of events may distinguish events due to apneas from events due to gross body movements. For each event we therefore calculated the total power modulation (TPM), comparing the 2 s before, to the 2 s after the  $M$  maximum:

$$TPM(t_M) = \frac{\langle TS \rangle_a - \langle TS \rangle_b}{\langle TS \rangle_a + \langle TS \rangle_b} \quad (37)$$

$$a = [t_M, t_M + 2s]; b = [t_M - 2s, t_M];$$

Presumably, the TPM feature has a small and often negative value for apnea events, and a high value (positive or negative) for gross body movement events.

The two quantifiers (36) and (37) are then used to train a support vector machine (SVM) as in the previous application. We refer to [115] for further details.

*Monitoring of the respiratory rate in infants between 2 and 6 months of age* is another application of respiratory rate detection. It is critical in infants because unprovoked respiratory arrest (for some reason, most often during deep sleep, the baby “forgets” to breathe) is the leading cause of SIDS, especially in infants between 2 and 6 months of age. As in the previous task of detecting respiratory arrests, particularly central apnea, we developed a reliable, automated, non-contact algorithm for real-time respiratory rate monitoring using a video camera. The settings for the present task are well defined, since the baby lies swaddled in a crib, and the camera is mounted above the crib. This allows for easy preliminary selection of a rectangular ROI that lies close to the frontal camera plane, covering the chest and abdomen, i.e., the places where the dominant respiratory movements (expansion and contraction) occur.

In the patent application [US20230270337A1] six methods are proposed for detecting the respiratory rate  $S(t)$  that may be used in different situations.

The total movement in the video is quantified as in the previous applications by the spectral optical flow algorithm GLORIA giving directly the rates of the six motions (20) in the plane  $V(t) = \{V_c(t) \mid c = 1, \dots, 6\}$ .

In some cases, the number of motion degrees can be reduced. For example, when the ROI is in a plane close to the frontal plane, the vectors (elements of the vector field corresponding to the motions caused by respiration) are decomposed mainly along the dilation and shear transformation axes, with only a small part of the respiration being projected onto the translation and possibly rotation axes. In such a case, both translations and rotations may be omitted due to the lack of quantification of the respiratory signal. In other cases, due to the presence of different objects in the video or different camera positions, only rotation may be omitted or a different set of quantifier motions may be chosen.

The features of the vector  $V(t)$  are analysed in a specific frequency interval (Breathing Frequency of Interest (BFOI), or the interval of frequencies to be used to look for respiratory activity [0.5, 1] Hz). After filtering, we created a new signal  $Vh(t)$  containing only periodicities between 0.5 and 1 Hz. The featured signal  $V(t)$  may be filtered in any manner. For example, using a band-pass filter (e.g., using the Fourier transform, Gabor wavelets, or a Hilbert-Huang transform (HH)). Then a respiratory rate detector  $S(t)$  is built. The  $S(t)$  outputs a time-dependent signal  $S_n(t)$  that represents the breathing of the subject where the distinct local maxima(s) above a certain threshold represent the times of inhaling and the local minima(s) under a certain threshold represent the times of exhaling.

The abbreviations for the six methods are: VFD, GSD, ICD, GMD, HFD, and SCD. Here we give a brief description of each of them.

**VFD** is a data-driven method, where the filtered signal is used directly. Therefore, the method directly uses the selected components of the movements' vector field between every two consecutive frames. The filtered signal represents a series of time vectors, and therefore, a complex time series is assumed. The selected complex time series is summarised (vector sum). Then, the obtained complex time series is divided into two time series, one for its real part and one for its imaginary part. The upper spline envelope over the maxima of the real-part time series represents the  $X$  coordinates of the proposed respiratory rate detector, corresponding to time. The upper spline envelope over the maxima of the imaginary-part time series represents the  $Y$  coordinates of the proposed respiratory rate detector corresponding to the magnitude. This results in a determination of  $VFD(t)$ , which means the time (i.e., the  $X$ -coordinates) and magnitude (i.e., the  $Y$ -coordinates) of the proposed respiratory rate detector  $S(t)$ .

The **GSD** is a model-driven method, based on dimensionality decrease, where the breath detector can be built from the situational model. One may use GSD when there is only one type of movement in the frame, specifically the motions caused by the subject's breathing. In such a case, the rest is noise due to the camera itself and the recording conditions. The GSD may also be used when there are several other movements in the frame, partially correlated with breathing, caused by different physical sources, and where a linear correlation exists between the various physical sources and respiration. Then, the GSD eliminates the linear correlations between different data dimensions. GSD uses principal component analysis (PCA) to eliminate linear correlations between different data dimensions, thereby reducing noise and enhancing the signal generated by breathing, and reducing the dimensions to a single one. By doing so, the resultant signal (data in principle mode) corresponds to the subject's inhalations and exhalations, corresponding to the local maxima/minima of the signal.

The **ICD** is a model-driven approach based on dimensionality reduction. The ICD method is used when there are several other movements in the frame, partially correlated with breathing, caused by different physical sources, and a nonlinear correlation or no correlation between the various physical sources and respiration (as opposed to the linear correlation discussed above regarding the GSD method). In the ICD, independent component analysis (ICA) is employed to eliminate the nonlinear correlations, yielding a set of statistically independent signals. This result is based on the idea that different physical sources generate the resulting signals. In some cases, it may not be clear which of the obtained set of signals represents respiration, and it is also possible that the

breathing signal is not separated as an independent physical source. Still, rather, different parts of it are involved in the separated signals. In such cases, the ICD may still be used for building the respiratory rate detector  $S(t)$  when additional information is available that can answer these questions. Otherwise, one may use the methods discussed below.

The **GMD** is a model-driven method, based on dimensionality reduction. The GMD method may be used when additional information is available that can address the questions discussed above regarding the ICD method. For the GMD method, the data is in the principal mode (after, for example, the use of the GSD method, discussed above). In the GMD method, the received signals are embedded in time to obtain the unknown variable and possible time dependencies. Then, the observed time series are converted to state vectors. A phase-space reconstruction procedure can then be used to verify the system order and reconstruct all dynamic system variables while preserving the system's properties. For each time series, the phase-space dimension and time lag are estimated, and the highest time lag, along with its corresponding phase-space dimension, is chosen. For all the time series obtained so far, the phase space of the uniformly sampled time-domain signal is reconstructed with the chosen couple: a time delay and embedding dimension. PCA may then be used for every newly obtained state vector to build a new system (e.g., resulting signals matrix after PCA) with removed time delay correlations (using PCA to put the resulting signals' matrix into a principal mode again, removing the linear correlations between the nonlinearities caused by time embedding). In some examples, the resulting signal matrix after PCA may still contain linear or nonlinear trends. In such cases, the typical trend may be removed by summing the time series cumulatively and detrending the results. This may result in a system noise that persists in the result. If so, the detrended signal may be decomposed into components using the HH, and the first HH component may be used as a resulting signal (thereby removing the HH residual).

The **HFD** is a model-driven, spectral contrast-based detector. In the HFD, the empirical decomposition-based detector's signal can be represented by the ratio between the sum of the second spectrum of the HH transform components of the signals  $V_h(t)$  and the sum of the second spectrum of the HH transform component of the signals  $V(t)$ .

**SCD** builds a model-driven, spectral contrast-based detector. In the method, to use the frequency (spectral) information, the time-dependent spectral content is calculated in each of the chosen group velocities using convolutions with Gabor filters, as described in the *Detection of respiratory arrests* section.

The respiratory rate detector  $S(t)$  may be constructed using two or more of the methods mentioned above. For example, the respiratory rate detector  $S(t)$  may be built (first) by using the first, second, third, and fourth methods (i.e., the VFD method, the GSD method, the ICD method, and the GMD method discussed above), and then (second) by using the fifth and sixth methods (i.e., the HFD method, and the SD and SCD methods discussed above). Preferably, the respiratory rate detector  $S(t)$  may be constructed (first) by using the fourth method (i.e., the GMD method), and then (second) by using the sixth method (i.e., the SD and SCD methods). Sometimes, more than one of the first, second, third, and fourth methods (discussed above) may be selected for use in constructing the respiratory rate detector  $S(t)$ . Additionally, in some cases, both the fifth and sixth methods can be chosen. The use of additional methods confirms the initial results and leverages some specific aspects of the current case. One or more artificial intelligence or machine learning methods can be employed to detect and recognise anomalies, thereby adjusting the system's response. The final step is to join some consecutive local maxima that are too close to each other in time.

For the examples shown in the Results section, we derived time series, representing the spatial transformations or group velocities: translational velocities along the two image axes. Here, we initially chose to omit rotation due to the lack of quantification of the respiratory signal. Furthermore, we construct the respiratory detector as follows:

We obtained the time-dependent spectral composition by averaging the time series over the five group velocities. We then filtered the resulting signal using an empirical decomposition with a stopping criterion for the last level that has at least  $2 * T$  maxima (where  $T$  is the recording time in

seconds). We then used the HFD to obtain  $S1(t)$  as the respiratory detector. The initial assumption is that the individual local maxima of  $S1(t)$  represent the respiratory times. Where by “separate local maxima” we assume a threshold  $\tau$  that connects some maxima to one if in time they are too close ( $\Delta t_{i,i+1} < \tau$ ) to each other. Therefore, we joined several detections of the same event.

### 2.7. Detection and Charge Estimation of Explosions

In the original publication [116] we have shown that three-dimensional scene can be reconstructed from images taken from multiple cameras situated in general positions and intersecting their fields of view. This reconstruction can be used to localize explosions and estimate its charge. Here we reproduce only the part of the methodology related to the use of optical flow.

To localize specific events in each of the camera's images, global motion reconstruction provided by the GLORIA algorithm is not sufficient. For the purpose we need the complete vector displacement, or velocity in this application, field. One possible approach is to apply the SOFIA algorithm, where equation (11) gives the reconstructed local and instantaneous velocity field  $\vec{v}(x, y, t) \equiv \{v_x(x, y, t), v_y(x, y, t)\}$ . We omit the scale parameter or sequence of scales for simplicity.

Explosions are detected as expansion events that can be characterized by the high positive divergence of the vector field. Because of the high velocities associated with such events, high speed cameras with 6000 frames per second were used. To avoid local fluctuations, we define a smoothed Gaussian spatial derivative of the vector field:

$$\begin{aligned} \nabla_a^\sigma \vec{v}(\vec{r}, t) &\equiv \iint \frac{\partial}{\partial r_a} G(\vec{r} - \vec{\rho}, \sigma) \vec{v}(\vec{\rho}, t) d^2 \vec{\rho} \\ \vec{r} &\equiv \{x, y\} \equiv \{r_1, r_2\} \\ G(\vec{r}, \sigma) &\equiv \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{|\vec{r}|^2}{\sigma^2}} \end{aligned} \quad (38)$$

The divergence of the vector field at the selected scale (our choice was  $\sigma$  of 30 pixels but the results were not sensitive to this parameter) is:

$$Q^\sigma(\vec{r}, t) \equiv \nabla_x^\sigma v_x(\vec{r}, t) + \nabla_y^\sigma v_y(\vec{r}, t) \quad (39)$$

From the quantity (39) we can localize the coordinates in the image plane and video sequence time (frame) of potential explosion events

$$\{\vec{r}_E, t_E\} = \operatorname{argmax}_{\vec{r}, t} (Q^\sigma(\vec{r}, t)) \quad (40)$$

The localization procedure is done simultaneously in all camera registrations and the position of the explosion in the three-dimensional scene is reconstructed from the generic formalism developed in [116].

Finally, as an overall estimation of the released energy by the explosion, the following expression was proposed:

$$W = \sum_{t=t_E}^{t_E+T} \iint |\vec{v}(\vec{r}, t)|^2 d^2 \vec{r} \quad (41)$$

This quantity is calculated for all camera registrations and added with the corresponding distance corrections. We selected  $T=100$  frames corresponding to  $1/60$  of a second, the time for a sound wave to cover slightly over five meters.

## 2.8. Object Tracking

Tracking of moving objects by using the global motion optical flow reconstruction method is introduced in detail in [133]. Although it can be applied to any group of transformations, our choice here is on the two translation rates and the dilatation (a global scale factor quantity) that are provided by the first three generators from Equation (20). We mark for clear interpretation the triplet  $\{A^1, A^2, A^3\}$  of reconstructed parameters in (19) as  $T_x^i$  and  $T_y^i$  for the translations and  $D^i$  for the dilatation, where  $i$  indicates which two consecutive frames  $\{i-1, i\}$  were used for the calculation. We restrict the current method to only these three transformations because we do not intend to rotate the region of interest (ROI) with the tracked object nor change the ratio between the ROI dimensions -  $L_x, L_y$ . In this way, our method is directly applicable to a situation where pan, tilt, and zoom (PTZ) hardware actuators are affecting the camera field of view that corresponds to the two translations (pan and tilt) and the dilatation (the zoom). Accordingly, we define the dynamic ROI with a triplet of values  $\{X_c^i, Y_c^i, L^i\}$  representing the coordinates of the ROI center and the length of the ROI diagonal  $L = \sqrt{L_x^2 + L_y^2}$ . Because of the fixed, constant ratio between the ROI dimensions, these three parameters uniquely define the ROI at each frame  $i$ .

In this notation, the ROI transformation driven by the translations and dilatation reconstructed parameters is:

$$\begin{aligned} X_c^i &= X_c^{i-1} + T_x^i; & Y_c^i &= Y_c^{i-1} + T_y^i \\ L^i &= L^{i-1} * (1 + D^i); \end{aligned} \quad (42)$$

Equations (42) define the ROI transition between from frame  $i-1$  to frame  $i$ . Note that in the size transformation of ROI, we have assumed that for infinitesimal dilatations  $e^D \cong 1 + D$ .

We have developed an extension of the single camera tracking algorithm to simultaneous multi-camera ROI tracking in [143]. In its simplest form, a linear model can describe the relationship between the tracking processes from  $N$  cameras:

$$ROI_a(k) = \sum_{b \neq a}^N W_{ab} * ROI_b(k) + R_a \quad (43)$$

In (43)  $a, b = \{1, \dots, N\}$  are the labels of the individual cameras.  $W_{ab}$  are  $3 \times 3$  transitional matrices and  $R_a$  are  $3 \times 1$  offset vectors;  $ROI \equiv \{X_c^i, Y_c^i, L^i\}$  is considered a vector as defined above. In the original work [143] a dynamic reinforcement algorithm based on quadratic cost-function minimization is proposed that can determine the  $W_{ab}, R_a$  interaction parameters from the tracking process. This way, the individual cameras start the tracking independently but, in the process, they begin to synchronize their ROIs. The linear model (43) has limited applications and we introduced non-linear interactions between the tracking algorithms. The details go beyond the scope of this review.

## 2.9. Image Stabilizing

The challenge of stabilizing image sequences affected by camera motion artefacts can be formulated as follows. Let the image sequence  $\{L^c(x, t + k\delta t)\}$ ;  $k = 0, 1, \dots, n$  contain an initial image for  $k = 0$  and the subsequent registrations that are affected or shifted by the motion artefacts. The objective of the methodology patented in [US 2022/0207657] is to build a filter that restores the sequence at any discrete index  $k > 0$  to the initial image that is conveniently chosen. To this end we recall Equation (2) and introduce an extra notation

$$L^c(x, t + k\delta t) = L^c(x - v^{(k)}(x), t + (k+1)\delta t) \equiv D_{v^{(k)}}\{L^c\} \quad (44)$$

Here  $D_v\{L^c\}$  is a short abbreviation for the vector diffeomorphism acting on the image. Note that here we have used the image optical flow transformation "in reverse". We use the optical flow

reconstruction algorithm, introduced in the first two subsections of the methods, to find the vector diffeomorphism  $v^k(x)$  that returns the current image to the previous one. Stabilizing the image sequence and removing motion artefacts due to camera motion involves reconstructing the corresponding vector field that connects the shifted images at all times to the initial one. To achieve this, we first stress that the application of two successive morphisms  $v(x)$  and  $g(x)$  is not equivalent to one with the sum of the two vector fields. More precisely, we need to “morph” the first vector field (shift its spatial arguments) by the second one:

$$D_g\{D_v\{L^c\}\} = D_{g+D_g\{v\}}\{L^c\} \quad (45)$$

Therefore, the resulting vector field generating the diffeomorphism from two successive vector diffeomorphisms is:

$$V(g, v) = g + D_g\{v\} \quad (46)$$

The above equation gives the group convolution law for vector diffeomorphisms. We apply (46) iteratively to reconstruct the global transformation between the initial image and any subsequent image of the sequence:

$$V_{k+1} = V_k + D_{V_k}\{v_k\} \quad (47)$$

Here,  $v_k$  is the infinitesimal vector-field transformation connecting the images  $L^c(x, t + k\delta t)$  and  $L^c(x, t + (k + 1)\delta t)$ . The resulting aggregated vector field  $V_k(x)$  connects the  $k$ -th image to the original member  $L^c(x, t)$  of the video sequence. Note again that also in (47) the diffeomorphisms transform the sequence members in the reverse direction, from the current image to the initial.

We define therefore the stabilized  $k$ -th image as

$$SL^c(x, t + k\delta t) \equiv D_{V_k}L^c(x, t + k\delta t) \quad (48)$$

The above Equation (48) is the required filter that “recovers” the shifted images and transforms them closest to the initial one. The latter can be chosen arbitrarily. Depending on the application, it can be updated at any fixed number  $n$  of images, or updated when some appropriate condition is met, for example when the aggregated vector field  $V_k(x)$  exceeds certain norm and the stabilization procedure becomes unfeasible.

### 3. Results

In the above technique, we can use either the local OF reconstruction approach SOFIA (11) or the global motion OF reconstruction GLORIA (14). In the first case, we attempt to filter all changes due to movements in the image sequence. Perhaps more flexible as well as computationally faster is the second option. GLORIA algorithm allows selecting a subset of transformations to filter out, leaving the rest of the movements intact. In the case of oscillatory movements of a camera, we can choose to filter only one or both of the translational movements. Rotational, dilatational and other displacements will be still present in the video sequence as they may be part of the intended observation content.

In the next subsections, we show summaries of the main results reported in our works ordered by the applications presented in the previous section.

#### 3.1. Spectral Optical Flow Iterative Algorithm (SOFIA)

The accuracy of optical reconstruction has been evaluated for multiple images, transformation fields and reconstruction parameters [6]. Our method, applied iteratively with sequence of scales [1, 2, 4, 8, 16] significantly outperforms the standard Matlab® Horn-Schunk ‘opticalFlowHS’ routine with default parameters of smoothness: 1, Maximal Iterations: 10 and minimal velocity 0. The average reconstruction error, tested for 8 images and 20 random vector deformation fields of average magnitude 0.5 pixels and spatially smoothed to 32 pixels was 2.5%. Our results show also that the

reconstruction precision depends on the number of iterations going from large to fine scales. Table 1 gives the average reconstruction error as function of the iteration scale.

**Table 1.** Average reconstruction error as a function of the iteration scales used.

Scales [pixels]	Error %
[16]	30
[8, 16]	10
[8,1]	5
[16, 8, 4, 2]	3
[16, 8, 4, 2, 1]	2.5

We found also that the spectral content of the image can influence the accuracy of the OF reconstruction. Our method is intrinsically multi-spectral; images with low spectral dispersion (like monochromatic ones) give higher error than images with balanced spectral content. In addition, images with higher spatial wavelengths give better OF reconstruction accuracy than images containing more short distance details (textures). For a quantified version of the above statements we refer to the original work [6].

### 3.2. Global Lie-Algebra Optical Flow Reconstruction Algorithm (GLORIA)

Following the results from the validation tests described in [8], GLORIA reconstruction applied with the transformations (20) gives accuracy depending on the magnitude of the transformations. In Table 2 we show the average errors for the corresponding group parameters.

**Table 2.** Average reconstruction error as a function of the iteration scales used. The relative coefficient differences (in %) were averaged for 10 images and 40 randomly generated transformation vectors (N=400), for each magnitude value. The error in % in the second column is the rounded average over all 6 transformations.

Magnitude [pxls]/Transformation	1	2	3	4	5	6
Reconstruction error [%]	2	5	6	7	7	8

### 3.3. Detection of Convulsive Epileptic Seizures

Here we present some validation results that are from offline application of our seizure detection algorithm. We first note that there are several different instances of the detector and the specific details are reported in the corresponding original works. The basic component is however, the use of optical flow motion parameters reconstruction and a subsequent spectral filtering.

In the seminal work [36] we have analyzed the performance of the detector in 93 convulsive seizures recorded from 50 patients in our long-term monitoring unit. We show that for a suitable selection of the detection threshold, a sensitivity of 95% and a false positive (FP) rate of less than one FP per 24 hours is achievable.

Automated video-based detection of nocturnal convulsive seizures was later investigated in our residential care facility [35]. From 50 convulsive seizures, all were detected (100%) sensitivity and the FP rate was 0.78 per night. The detection delay in 78% of the cases was less than 10 seconds; maximal delay was of 40 seconds in one case. There were also other types of epileptic seizures registered in the study; the detector was less sensitive to motor events of non-convulsive patterns.

Detection and alerting for convulsive seizures in children were conducted in [34]. The dataset included 1661 full recorded nights of 22 children (13 male) with a median age of 9 years (range of 3-17 years). The video detection algorithm was able to detect 118 of 125 convulsive seizures, overall sensitivity 94%. The total FP detections were 81, rate 0.048 per night.

The adaptive paradigm proposed in [37,38] was tested on one patient exhibiting frequent tonic-clonic convulsive seizures. The total observation time was 230 days with 228 events detected by the system. This case study showed that with the default parameter settings the specificity, the

percentage of true alarms was 70% corresponding to an average of 1 FP per 2.6 days. After applying parameter reinforcement optimization, the specificity was elevated to 93% corresponding to an average of 1 FP per 18 days. Unfortunately, no “ground truth” tracking of all possible seizures was available, the patient was in residential setting and no continuous video monitoring was installed. Therefore, we cannot report on the sensitivity for this cases study.

### 3.4. Forecasting (PGES)

In [41] we found that in accordance with results from a computational model, clinical clonic seizures exhibit an exponential decrease of the convulsion frequency or, equivalently, exponential increase of the inter-clonic intervals (ICI). We found also that there is a correlation between the terminal ICI and the duration of a post ictal suppression, PGES phase. The relation between the two was estimated from analyzing 48 convulsive seizures 37 of which resulted in PGES phase. The association measure is the amount of explained variation between two time series and is defined as:

$$h^2(T_{PGES}, ICI_{terminal}) \equiv 1 - \frac{\text{var}(T_{PGES} | ICI_{terminal})}{\text{var}(T_{PGES})} \quad (49)$$

It is clear from Equation (49) that if the conditional variation between the two quantities is zero, meaning that one is an exact function of the other, the index is one. If the two quantities are independent, the conditional variation is equal to the total one and the index is zero. Note that the index (49) is asymmetric to its arguments. The value of this index for our sample series was 0.41. It is not a large value but it is statistically significant. The statistical significance of the index (49) can be estimated by taking a number (100 or more) of random permutations of the time stamps in one of the signals and calculating (49) in each of them to establish the probability  $p$  of obtaining the specific association value or higher by chance. In all our reported results we have at least  $p < 0.05$ .

To automate the process of estimating the increase rate of the ICE from the OF analysis, we applied the technique in subsection 2.4 to 33 video sequences [40]. We found that the association indexes (49) between the manual and automated rate estimates are  $h^2(automated, manual) = 0.87$  and  $h^2(manual, automated) = 0.74$ .

The efficient automated procedure allowed for further investigation of the relations between the PGES duration and the exit ICI in convulsive seizures. In [42] 48 cases of convulsive seizures with PGES and 27 without PGES were analyzed. An SVM classifier using the exit ICI and the seizure duration was constructed and after 50-fold training-performance repetitions, we reached a mean accuracy of 99.7%, mean sensitivity of 99.0% and mean specificity of 100%.

### 3.5. Detection of Falls

In the original work [87] we used two datasets for the development and testing of the fall detection algorithm; the publicly available Le2i fall detection database [60] and the SEIN fall database, a video database of recordings of genuine falls from people with epilepsy, collected at our center. The Le2i database contains 221 videos simulated by actors, with falls in all directions, various normal activities and challenges such as variable illumination and occlusions. Some of the videos were without audio track and were excluded, leaving 190 video fragments used for training and evaluation. The overall results from classifiers using only the video information (features  $\{A, V, D\}$ , see Section 2.5) and the full video and audio features  $\{A, V, D, S\}$  are summarized in the Table 3 below.

**Table 3.** Fall detection performance results for the Le2i test set. Results from using the full feature set and from using only video features are shown. Specificity (SPEC) are given for three working points on the ROC curves chosen according to their sensitivity (SENS) values. ROC AUC are the receiver operating characteristic area under the curve.

DATA \ RES	ROC AUC	SPEC @ 100% SENS	SPEC @ 90% SENS	SPEC @ 80% SENS
Video & audio	0.957	0.818	0.919	0.945
Video only	0.947	0.799	0.896	0.923

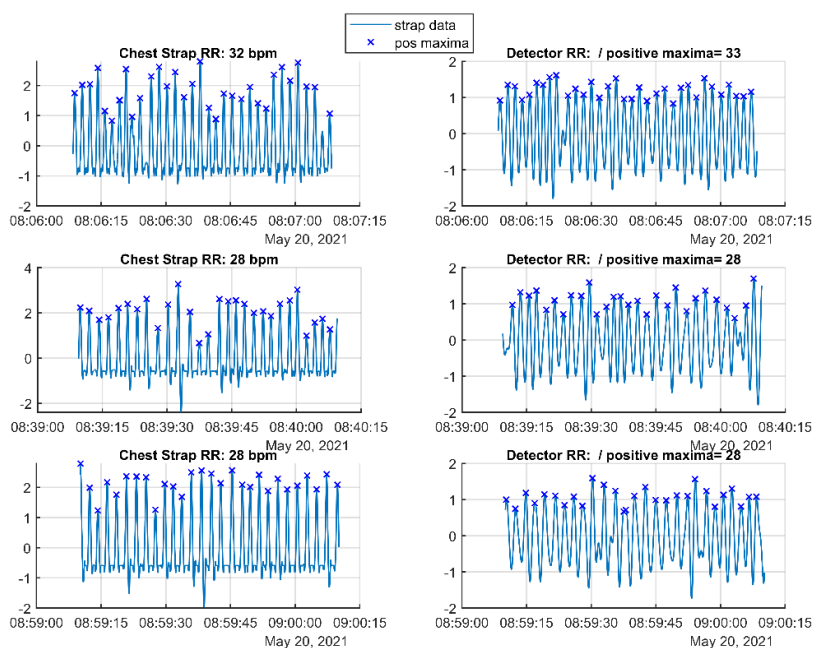
Recently we applied for the full Le2i dataset a more advanced machine learning paradigm in [93] using only video data, but considering all six global movement parameters instead of only the vertical translational component, and we achieved a ROC AUC of 0.98.

### 3.6. Detection of Respiratory Arrests, Apnea

The results reported in [115] suggest that the position of the camera largely influences the detector performance. Sensitivity varies from 80% (worst position) to 100% (best position) and the average from all positions was 83%. The corresponding false positive rates (events per hour) were between 3.28 and 1.09, the average for all the positions: 2.17.

In addition, we tested also an early integration between the camera signals. In the averaged spectrum of the OF from Equation (22), third line, traces reconstructed simultaneously from all cameras were included. The sensitivity was 92.9% and the false positive rate 1.64 events per hour. These numbers are in-between the best and worst camera positions but better than the averaged single-camera performance. Such result is especially interesting in cases when the best camera position is unknown or the position of the patient may change during the observation.

To show the results of *monitoring of the respiratory rate in infants between 2 and 6 months of age*, we compared the proposed method with a ground truth, namely “Chest Strap” - a recognized (contact) method for detecting the rhythm of breathing. Figure 2 shows *Chest Strap RR* (respiratory rate) and *Detector RR* readings on the same one-minute segments of three infants.



**Figure 2.** Comparison of *Chest Strap RR* and *Detector RR* readings for respiratory rate (RR) calculated on the same one-minute segment for the monitored three different infants. The left column shows the *Chest Strap RR* readings (ground truth), and the right column shows the *Detector RR* readings.

The mean respiratory rhythms for all of the examined infants are shown in Table 4

**Table 4.** The results of the two measurements of the mean respiratory rhythm in 7 babies aged between 3 and 5 months. The second and third column present respiratory cycles per minute.

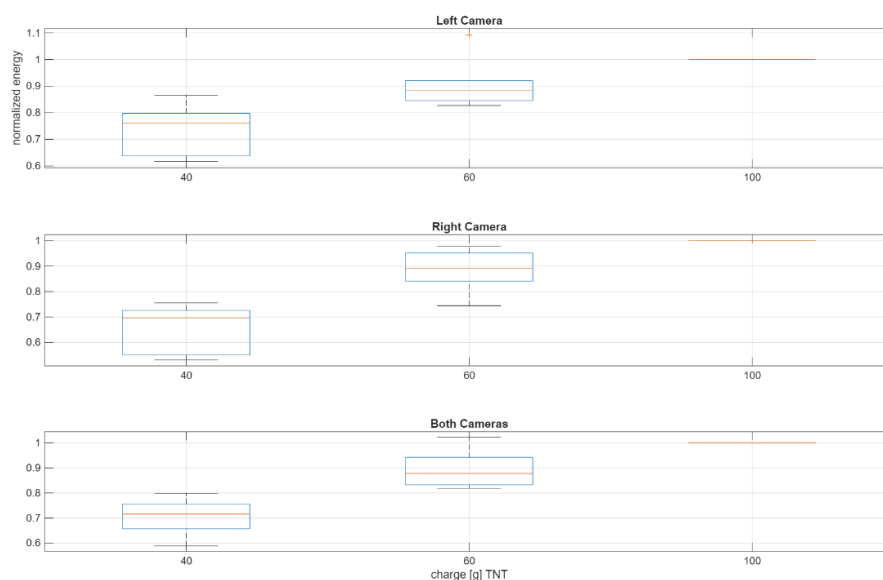
Video file index (.mp4)	A = Detector RR	B = Chest Strap RR	(A-B)
01	45	43	2
02	39	37	2
54	47	46	1
55	38	37	1
56	48	47	1
58	48	45	3
59	45	44	1

The duration of the measurements (movies included in Table 4) is between 2 and 6 hours, in which the sleep phases alternate with the awake phases of the babies.

### 3.7. Detection and Charge Estimation of Explosions

In the article [116] explosions of three different charges 40,60 and 100 gram of TNT were performed at six locations. The spatial reconstruction and subsequent charge estimations were done by registrations with two cameras installed on separate locations at approximately 10 meters from the explosions. The reconstructed 3D coordinates from the OF localization in each camera were within 200mm of the actual explosion locations. The maximal relative error was therefore  $0.2/10=0.02$ , or 2%.

Charge estimation was done for each camera separately and also by combining the energy estimates (41) of both cameras. In the original work we presented the raw estimates, here we also normalized all energy estimates to the corresponding ones from the largest charge (here with 100-gram TNT) for each explosion location in order to cancel the dependence on the distance to the camera.



**Figure 3.** The distributions over the six explosion locations of the normalized (to the charge of 100 g TNT) energy estimates registered from the left (upper plot), right (middle plot) and both (lower plot) cameras according to

the test charge (horizontal axes in gram TNT). The boxplots show the average (red lines) normalized energy, the 25 and 75 percentiles (box tops and bottoms) and the 10 and 90 percentiles (the whiskers). Red stars are the outliers.

From the figure, we see that the left camera gives better separation between the registered charges than the right one, while the combined estimate from both cameras interpolates the results.

### 3.8. Object Tracking

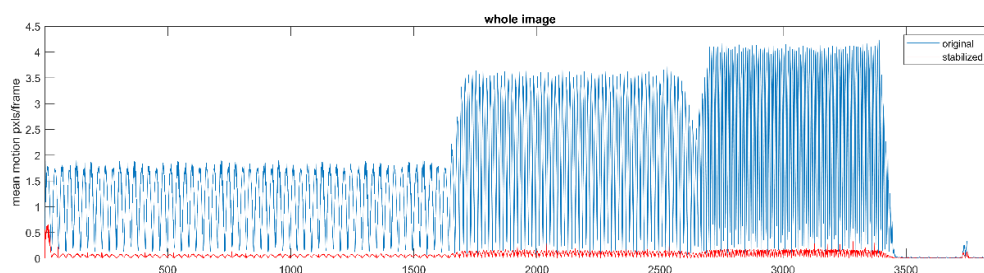
The tracking algorithm based on Equation (42) was validated in [133] on both synthetic motion sequences and real-world registrations. In the first case, we have a ground truth for the actual displacement parameters and for the second, operator tracking gave the “gold standard”. In all cases the overall quality of automated tracking at every time sample (or frame number)  $t$  is evaluated by the total deviation of the ROI coordinates  $\Delta_{total}(t) \equiv \sqrt{\Delta X_c(t)^2 + \Delta Y_c(t)^2 + \Delta L_x(t)^2 + \Delta L_y(t)^2}$ .

In the tests with synthetic images (Gaussian blob moving with 2 pixels per frame change in the x-direction and 1 pixel per frame change in the y-direction) the deviation was 0.05 pixels for both directions resulting in a relative error of 2.5% and 5% for the (x, y) directions correspondingly. Dilatations were tracked with 10% relative deviation.

In the follow-up work [143] the effect of reinforcement between the tracking algorithms of two cameras was studied. Fifty-one videos were generated. The total deviation in both cameras was calculated and averaged over all frames. The linear fusion model showed marginal improvement, the deviation was reduced by less than 4%. Non-linear interaction between the tracking sequences resulted on average in 30% reduction of the deviation between the tracked and target ROI. We also investigated the influence of object speed on the effectiveness of the non-linear reinforcement. The effectiveness decreased with the increase of object velocity, the approach increased however significantly the accuracy of tracking of objects moving slower than 0.3 pixels/frame.

### 3.9. Image Stabilizing

The methodology published in the patent “US 2022/0207657 GLOBAL MOVEMENT IMAGE STABILIZATION SYSTEMS AND METHODS” was tested on multiple scenarios of moving cameras, moving objects or both. The extended analysis and validation of the method will be reported in a separate work. Here we present the result from a simple test where the camera was subject to oscillatory movements and the stabilizing algorithm was based on the global movement OF reconstruction GLORIA involving translations, rotation, dilatation and shear transformations. In Figure 4 we show the motion content of a video sequence, measured by the pixel-level OF reconstruction method SOFIA, before and after the stabilizing process.



**Figure 4.** The effect of stabilizing of a video sequence affected by oscillatory movements. A sequence of three different frequencies and amplitudes is used. The blue trace is the mean OF frame-to-frame displacement in pixels of the original sequence. The red trace is the mean displacement of the stabilized image. The horizontal axis represents the frame number.

The test demonstrates that the stabilizing algorithm compensates more than 95% of the motion related OF amplitude.

## 4. Discussion

Here we discuss the general concepts as well as some specific issues related to the methods and applications reviewed in this work. We also outline some limitations of our approaches and accordingly, speculate about possible extensions and future research.

Most of the challenges where we applied the optical flow concept relate to detection and awareness of events such as motor epileptic seizures, falls, apnea. In the case of post-ictal electrographic suppression prediction, the method can be used for both real-time alerting and off-line diagnosis of cases with higher risk of PGES. We note however that in general the task of detection events in real-time is related but not equivalent to classification of signals. The essential difference is in the requirement to recognize the event as soon as possible without possessing the data from the whole duration of the event. Classification of off-line data can be important for diagnostic purposes, but for real-time detection of convulsive seizures for example, reaction times within 5-10 seconds achievable with our technique [35], can be critical for avoiding injuries or complications. The two objectives, classification and alerting, can be part of one system in the context of adaptive approaches involving machine-learning paradigms. In [37] we have used off-line cluster-based classification of already detected or suspected events [38] as part of unsupervised reinforcement learning procedure for fine-tuning the on-line detector. The assumption is that the OF signal during the total duration of the convulsive seizure can provide reliable discrimination between the real seizure detections from the false ones. Therefore, detector parameters dynamically adapt to the classification of the previous detections used as training sets. This approach was applied and tested only for the seizure detection but in the future, it may be used in other adaptive detectors. In this context we realize also that for some alerting applications, machine-learning approaches can be difficult to develop and their advantages can be disputable. Falls for example, happen due to a broad variety of factors and unsupervised learning approaches may not be effective. Providing training sets for all of them on the other hand, can be a challenge as well. Validating and labeling cases is also a time-consuming process and in addition depends on the skills of the qualified observers. In such applications, universal model-based algorithm may provide a feasible alternative. Our guiding principle is the “hybrid” approach, using as much as possible model based “backbone” algorithms such as the computational model induced post-ictal suppression prediction in [41]. The refinement of the detectors or predictors can be further achieved by machine learning paradigms.

In the context of the previous comments, state classification may provide predictive information about forthcoming adverse events. We have explored such possibilities in the cases of PGES by relating the convulsive movements dynamics to post-seizure suppression of the brain activity. Another example is the observation that respiratory irregularities may be prodromal for the catastrophic events of SIDS. We have analyzed also possibilities for short-term anticipation of epileptic seizures, the results are promising but more statistical evidence has to be collected.

Both SOFIA and GLORIA algorithms provide early multi-channel data fusion. As seen from the Equations (2), (3), (4) and (15), the velocity, or displacement field to be reconstructed is common for all the spectral channels, or colors in case of traditional RGB camera. Additional sensor modalities such as thermal (contrast) imaging, depth detectors, radars or simply broader array of spectral sensors can be included. The intrinsic multichannel nature of our algorithms decreases the level of degeneracy of the inverse problem. OF reconstruction is in general, and especially in case of using single channel intensity images, an under determined problem as the local velocities in the directions of constant intensity can be arbitrary. This is less likely to occur in multichannel images and therefore early data fusion is advantageous for obtaining robust solution.

The image sequence-based reconstruction of global movements further allows for early integration, or fusion, of multi-camera registrations. Because the spatial information is largely truncated, time series from the cameras can be analyzed simultaneously as was shown in [115] in the example of respiratory arrest detection. Signal fusion can be also done in later processing stages, as is the case with explosion charge estimates [116]. The synergy between OF algorithms running on a set of cameras can be achieved as a dynamic reinforcement process, as shown in the application of

tracking objects [143]. Sensor fusion paradigms can be also advantageous for the rest of the applications considered here and these may be subjects for further developments.

As described in the methods section, SOFIA is an iterative multi-scale algorithm. This means that we can control the levels of detail that we want to obtain in the solution. However, how do we choose these levels? In the current stage of applying the method, we have rigidly selected the sequences of scales according to the expected or assumed levels of detail that will be relevant for the specific analysis. In a more flexible and assumption free implementation, levels of detail may be possible to infer from the dynamic content of the video sequences. A simple approach will be to start at a coarse scale and then test whether the reconstructed displacement vector field “explains” sufficiently the changes in the frames. If not, a finer scale reconstruction will follow. We will address this extension in a future work.

Except for the part dedicated to image stabilization, all the applications here are assuming a static (or PTZ controlled in the case of object tracking module) camera observing scenes or objects. Optical flow derived algorithms can be extended to mobile cameras. The separation between the camera movement and the displacement of the registered objects will be subject to future investigations. One particular setup that can be of a direct benefit for the detection and alerting of convulsive epileptic seizures and of falls is the use of “egocentric” camera. The last can be the inbuilt camera of a smartphone, avoiding this way possible inconveniences associated with dedicated wearables. We believe that especially for detection of motor seizures, the direct application of the same algorithms used with static camera are applicable. The global movement reconstruction GLORIA may even be more effective in this setup as the whole scene will follow the convulsive movements of the patient. A test trial with wireless camera will be attempted in nearest future.

Our last remark concerns the issue of scalability of any system dedicated to real-world operation. Our approach, as stated in the Introduction and illustrated on Figure 1, allows for a common universal OF module linked to modular additions of various detectors. This feature distinguishes our paradigm from the variety of task-specific detectors that would require a separate processing implementation for each individual class of events. The last may be feasible only for small-scale applications like home use. In a typical care center, however, the number of residents that may need safety monitoring can be of order of 100 or more. It may be possible, but sometimes economically not realistic, that for each person a complete system will be installed. In addition, a video network supporting that many cameras (in some cases more than one camera per resident may be optimal) will be extremely loaded. Given that the OF reconstruction is the most computationally demanding part of the processing and that it is common for all detectors, a distributed system of smart cameras each with an uploaded GLORIA algorithm can provide the data for all the detectors running on centralized platform connected to observation stations. Indeed, OF signals are just six time-series per camera of relatively low sample rates (25-30 samples per second) and can easily be distributed to central processing servers over low-bandwidth network where the computationally light algorithms can run in parallel. We are considering these options within a pending institutional implementation phase.

## 5. Conclusions

The paradigm of optical flow reconstruction on a variety of levels, from fine scale pixel-level details to global movements, can be a common processing module providing data for a variety of video-based remote detectors. The detectors can be implemented separately, concurrently or working synchronously in parallel to selectively identify and alert for hazardous situations. Off-line implementations can be used for dedicated diagnostic or forensic algorithms. Global movement reconstruction can serve at the same time as input for automated tracking and image stabilizing algorithms. The major computational complexity, the OF inverse problem solution, is therefore centrally addressed providing thus significant reduction of subsequent processing resources.

## 6. Patents

1. Karpuzov S, Kalitzin S., Petkov A, Ilieva S, Petkov G, **METHOD AND SYSTEM FOR OBJECTS TRACKING IN VIDEO SEQUENCES**  
<https://patentscope.wipo.int/search/en/WO2025085981>
2. Petkov, G., Fornell, P., Ristic, B. and Trujillo, I., HB Innovations Inc, 2023. **System and method for video detection of breathing rates**. U.S. Patent Application 17/682,645.  
<https://patents.google.com/patent/US20230270337A1/en>
3. Petkov G, Kalitzin S, Fornell P. **Global movement image stabilisation systems and methods** [US PATENT US20220207657A1/US11494881B2 citations (17)/(5)]. Available from:  
<https://patents.google.com/patent/US11494881B2>

**Author Contributions:** Conceptualization, S.K.; methodology, S.K., S.B.K. and G.P.; software, S.B.K, G.P. and S.K.; data S.B.K.; validation, S.B.K, G.P.; writing—original draft preparation, S.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research is part of the GATE project funded by the Horizon 2020 WIDESPREAD-2018-2020 TEAMING Phase 2 programme under grant agreement no. 857155, the programme “Research, Innovation and Digitalization for Smart Transformation” 2021-2027 (PRIDST) under grant agreement no. BG16RFPR002-1.014-0010-C01. Stiliyan Kalitzin is partially funded by “Anna Teding van Berkhout Stichting”, Program 35401, Remote Detection of Motor Paroxysms (REDEMP).

**Institutional Review Board Statement:** Not applicable”

**Informed Consent Statement:** Not applicable

**Conflicts of Interest:** The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## Abbreviations

The following abbreviations are used in this manuscript:

OF	Optical Flow
SVM	Support Vector Machine
CNN	Convolutional Neural Network
ROI	Region Of Interest
PTZ	Pen, Tilt, Zoom
SUDEP	Sudden Unexpected Death in Epilepsy
PGES	Post-ictal Generalized Electrographic Suppression
FP	False Positive
ICI	Inter-Clonic Interval
TNT	Tri Nitro Toluene

## References

1. Beauchemin, S.S.; Barron, J.L. The computation of optical flow. *ACM Computing Surveys (CSUR)* **1995**, *27*(3), 433-466. doi:10.1145/212094.212141
2. Horn, B.K.P.; Schunck, B.G. Determining optical flow. *Artif Intell* **1981**, *vol. 17*, 1–3, 185–203. doi: 10.1016/0004-3702(81)90024-2.
3. Niessen, W.J.; Duncan, J.S.; Florack, L.M.J.; ter Haar Romeny, B.M.; Viergever, M.A. Spatiotemporal operators and optic flow. *Physics-Based Modeling in Computer Vision, IEEE Computer Society Press* **1995**, *7*.
4. Niessen, W.J.; Maas, R. Multiscale optic flow and stereo. In: *Computational Imaging and Vision*. Sporringer, J., Nielsen, M., Florack, L., Johansen, P., Eds.; Kluwer Academic Publishers, 1997, 31-42.
5. Maas, R.; ter Haar Romeny, B.M.; Viergever, M.A. A multiscale Taylor series approach to optic flow and stereo: a generalization of optic flow under the aperture. In: *Scale-Space Theories in Computer Vision*. Nielsen, M., Johansen, P., Fogh Olsen, O., Weickert, J., Eds.; Springer, 1999, vol. 1682, 519-524.

6. Kalitzin, S.; Geertsema, E.; Petkov, G. (2018), Scale-iterative optical flow reconstruction from multi-channel image sequences. In: *Frontiers of Artificial Intelligence and Applications*, Petkov, N., Strisciuglio, N., Travieso-Gonzalez, C., Eds.; IOS Press, Amsterdam, 2018, Vol 310, *Application of Intelligent Systems*, 302-314, DOI:10.3233/978-1-61499-929-4-302
7. Florack, L.M.J.; Nielsen, M.; Niessen, W.J. *The intrinsic structure of optic flow incorporating measurement duality. International Journal of Computer Vision* **1998**, 27(3), 24 .
8. Kalitzin, S.; Geertsema, E.; Petkov, G. (2018), Optical flow group-parameter reconstruction from multi-channel image sequences. In: *Frontiers of Artificial Intelligence and Applications*, Petkov, N., Strisciuglio, N., Travieso-Gonzalez, C., Eds.; IOS Press, Amsterdam, 2018, vol 310, *Application of Intelligent Systems*, 290 – 301, DOI:10.3233/978-1-61499-929-4-290
9. Sander, J.W. Some aspects of prognosis in the epilepsies: a review. *Epilepsia*. **1993**, 34(6), 1007-1016.
10. Blume W.T.; Luders, H.O.; Mizrahi, E.; Tassinari, C.; van Emde Boas, C.W.; J. Engel Jr., J. Glossary of descriptive terminology for ictal semiology: Report of the ILAE task force on classification and terminology, *Epilepsia*, **2001**, vol. 42, 1212–1218.
11. Karayiannis, N.B.; Mukherjee, A.; Glover, J.R.; Ktonas, P.Y.; Frost, J.D.; Hrachovy Jr., R. A.; Mizrahi, E.M. Detection of pseudosinusoidal epileptic seizure segments in the neonatal EEG by cascading a rule-based algorithm with a neural network, *IEEE Trans. Biomed. Eng.*, **2006**, vol. 53, no. 4, 633–641.
12. Becq, G.; Bonnet, S.; Minotti, L.; Antonakios, M.; Guillemaud, R.; Kahane, P. Classification of epileptic motor manifestations using inertial and magnetic sensors, *Comput. Biol. Med.*, **2011**, vol. 41, 46–55.
13. Surges R.; Sander J.W. Sudden unexpected death in epilepsy: mechanisms, prevalence, and prevention. *Curr Opin Neurol* **2012**, 25: 201–7.
14. Ryvlin P.; Nashef L.; Lhatoo S.D.; Bateman L.M.; Bird J.; Bleasel A. et al. Incidence and mechanisms of cardiorespiratory arrests in epilepsy monitoring units (MORTEMUS): a retrospective study. *Lancet Neurol* **2013**, 12, 966–77.
15. Van de Vel, A.; Cuppens, K.; Bonroy, B.; Milosevic, M.; Jansen, K.; Van Huffel, S.; Vanrumste, B.; Lagae, L.; Ceulemans, B. Non-EEG seizure-detection systems and potential SUDEP prevention: state of the art, *Seizure*, **2013**, 22, 345–355, <http://dx.doi.org/10.1016/j.seizure.2013.02.012>.
16. Saab M.E.; Gotman, J. A system to detect the onset of epileptic seizures in scalp EEG, *Clin. Neurophysiol.*, **2005**, vol. 116, 427–442.
17. Pauri, F.; Pierelli, F.; Chatrian, G.E.; Erdly, W.W. Long-term EEG video- audio monitoring: computer detection of focal EEG seizure patterns. *Electroencephalogr. Clin. Neurophysiol.* **1992**, vol. 82, 1–9.
18. Gotman, J. Automatic recognition of epileptic seizures in the EEG. *Electroencephalogr. Clin. Neurophysiol.* **1982**, vol. 54, 530–540.
19. Salinsky, M.C. A practical analysis of computer based seizure detection during continuous video-EEG monitoring, *Electroencephalogr. Clin. Neurophysiol.* **1997**, vol. 103, 445–449.
20. Schulc, E.; Unterberger, I.; Saboor, S.; Hilbe, J.; Ertl, M.; Ammenwerth, E.; Trinka, E.; Them, C. Measurement and quantification of generalized tonic-clonic seizures in epilepsy patients by means of accelerometry— An explorative study, *Epilepsy Res.* **2011**, vol. 95, 173–183.
21. Kramer, U.; Kipervasser, S.; Shlitner, A.; Kuzniecky, R. A novel portable seizure detection alarm system: preliminary results, *J. Clin. Neurophysiol.* **2011**, vol. 28, 36–38.
22. Lockman, J.; Fisher, R.S.; Olson, D.M. Detection of seizurelike movements using a wrist accelerometer. *Epilepsy Behav.* **2011**, vol. 20, 638–641.
23. van Andel, J.; Thijs, R.D.; de Weerd, A. et al. Non-EEG based ambulatory seizure detection designed for home use: what is available and how will it influence epilepsy care? *Epilepsy Behav.* **2016**, 57, 82–9.
24. Arends, J.; Thijs, R.D.; Gutter, T. et al. Multimodal nocturnal seizure detection in a residential care setting: a long-term prospective trial. *Neurology*. **2018**, 91, e2010–9.
25. Narechania, A.P.; Garic, I.I.; Sen-Gupta, I. et al. Assessment of a quasi-piezoelectric mattress monitor as a detection system for generalized convulsions. *Epilepsy Behav.* **2013**, 28, 172–6.
26. Van Poppel, K.; Fulton, S.P.; McGregor, A. et al. Prospective study of the Emfit movement monitor. *J Child Neurol.* **2013**, 28, 1434–6.

27. Cuppens, K.; Lagae, L.; Ceulemans, B.; Van Huffel, S.; Vanrumste, B. Automatic video detection of body movement during sleep based on optical flow in pediatric patients with epilepsy, *Med. Biol. Eng. Comput.* **2010**, vol. 48, 923–931.
28. Karayiannis, N.B.; Tao, G.; Frost Jr., J.D.; Wise, M.S.; Hrachovy, R.A.; Mizrahi, E.M. Automated detection of videotaped neonatal seizures based on motion segmentation methods, *Clin. Neurophysiol.* **2006**, vol. 117, 1585–1594.
29. Karayiannis, N.B.; Xiong, Y.; Tao, G.; Frost Jr., J.D.; Wise, M.S.; Hrachovy, R.A.; Mizrahi, E.M. Automated detection of videotaped neonatal seizures of epileptic origin. *Epilepsia*, **2006**, vol. 47, 966–980.
30. Karayiannis, N.B.; Tao, G.; Varughese, B.; Frost Jr., J.D.; Wise, M.S.; Mizrahi, E.M. Discrete optical flow estimation methods and their application in the extraction of motion strength signals from video recordings of neonatal seizures. in *Proc. 26th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2004**, vol. 3, 1718–1721.
31. Karayiannis, N.B.; Tao, G.; Xiong, Y.; Sami, A.; Varughese, B.; Frost Jr., J.D.; Wise, M.S.; Mizrahi, E.M. Computerized motion analysis of videotaped neonatal seizures of epileptic origin. *Epilepsia* **2005**, vol. 46, 901–917.
32. Chen, L.; Yang, X.; Liu, Y.; Zeng, D.; Tang, Y.; Yan, B.; Lin, X.; Liu, L.; Xu, H.; Zhou, D. Quantitative and trajectory analysis of movement trajectories in supplementary motor area seizures of frontal lobe epilepsy. *Epilepsy Behav.* **2009**, vol. 14, 344–353.
33. Li, Z.; Martins da Silva, A.; Cunha, J.P. Movement quantification in epileptic seizures: a new approach to video-EEG analysis. *IEEE Trans. Biomed. Eng.* **2002**, vol. 49, no. 6, 565–573.
34. van Westrhenen, A.; Petkov, G.; Kalitzin, S.N.; Lazeron, R.H.C.; Thijs, R.D. Automated video-based detection of nocturnal motor seizures in children. *Epilepsia* **2020**, 61(S1), S36–S40 DOI: 10.1111/epi.16504
35. Geertsema, E.; Thijs, R.D.; Gutter, T.; Vledder, B.; Arends, J.B.; Leijten, F.S.; Visser, G.H.; Kalitzin S.N. Automated video-based detection of nocturnal convulsive seizures in a residential care setting. *Epilepsia* **2018**, vol. 59, (S1), 53–60, DOI: 10.1111/epi.14050,
36. Kalitzin, S.; Petkov, G.; Velis, D.; Vledder B.; Lopes da Silva, F. Automatic Segmentation of Episodes Containing Epileptic Clonic Seizures in Video Sequences. in *IEEE Transactions on Biomedical Engineering*, **2012**, vol. 59, no. 12, 3379–3385, doi: 10.1109/TBME.2012.2215609.
37. Kalitzin S. Adaptive Remote Sensing Paradigm for Real-Time Alerting of Convulsive Epileptic Seizures. *Sensors* **2023**, 23, 968. <https://doi.org/10.3390/s23020968>
38. Kalitzin S. Topological Reinforcement Adaptive Algorithm (TOREADA) Application to the Alerting of Convulsive Seizures and Validation with Monte Carlo Numerical Simulations. *Algorithms* **2024**, 17, 516. <https://doi.org/10.3390/a17110516>
39. Surges, R.; Strzelczyk, A.; Scott, C.A.; Walker, M.C.; Sander, J.W. Postictal generalized electroencephalographic suppression is associated with generalized seizures. *Epilepsy Behav* **2011**, 21, 271–4.
40. Kalitzin, S.N.; Bauer, P.R.; Lamberts, R.J.; Velis, D.N.; Thijs R.D.; Lopes Da Silva F.H. Automated Video Detection Of Epileptic Convulsion Slowing As A Precursor For Post-Seizure Neuronal Collapse. *International Journal of Neural Systems* **2016**, 26(8), 1650027, DOI: 10.1142/S0129065716500271
41. Bauer, P.R.; Thijs, R.D.; Lamberts, R.J.; Velis, D.N.; Visser, G.H.; Tolner, E.A.; Sander, J.W.; Lopes da Silva, F.H.; Kalitzin S.N. Dynamics of convulsive seizure termination and postictal generalized EEG suppression. *Brain* **2017**, Volume 140, Issue 3, 655–668, DOI: <https://doi.org/10.1093/brain/aww322>
42. van Beurden, A.W.; Petkov, G.H.; N. Kalitzin, S.N. Remote-sensor automated system for SUDEP (sudden unexplained death in epilepsy) forecast and alerting: analytic concepts and support from clinical data. In *Proceedings of the 2nd International Conference on Applications of Intelligent Systems (APPIS '19)*. ACM, New York, NY, USA, Article 2, 1–6 DOI: <https://doi.org/10.1145/3309772.3309774>
43. Rubenstein, L.Z. Falls in older people: epidemiology, risk factors and strategies for prevention. *Age Ageing* **2006**, 35, ii37–ii41. <https://doi.org/10.1093/ageing/afl084>.
44. Davis J.C. et al., Cost-effectiveness of falls prevention strategies for older adults: protocol for a living systematic review *BMJ Open*, **2024**, vol. 14, no. 11, e088536, doi: 10.1136/BMJOPEN-2024-088536.

45. European Public Health Association, "Falls in Older Adults in the EU: Factsheet." [Online]. Available on site: [https://eupha.org/repository/sections/ipsp/Factsheet\\_falls\\_in\\_older\\_adults\\_in\\_EU.pdf](https://eupha.org/repository/sections/ipsp/Factsheet_falls_in_older_adults_in_EU.pdf) (Accessed: Jun. 04, 2025).
46. Davis, J.C.; Robertson, M.C.; Ashe, M.C.; Liu-Ambrose, T.; Khan, K.M.; Marra, C.A. International comparison of cost of falls in older adults living in the community: a systematic review. *Osteoporosis International* **2010**, *21*:8, vol. 21, no. 8, 1295–1306, doi: 10.1007/S00198-009-1162-0.
47. Krumholz, A.; Hopp, J.; Falls give another reason for taking seizures to heart. *Neurology* **2008**, *70*, 1874–1875. <https://doi.org/10.1212/01.wnl.0000312285.73631.ff>.
48. Russell-Jones, D.L.; Shorvon, S.D. The frequency and consequences of head injury in epileptic seizures. *J. Neurol. Neurosurg. Psychiatry* **1989**, *52*, 659–662. <https://doi.org/10.1136/jnnp.52.5.659>.
49. Nait-Charif, H.; McKenna, S. 2004. Activity summarization and fall detection in a supportive home environment. *IEEE International Conference on Pattern Recognition*, 26-26 Aug, **2004**, 20–23. DOI: 10.1109/ICPR.2004.
50. Wang, X.; Ellul, J.; Azzopardi, G. Elderly Fall Detection Systems: A Literature Survey. *Frontiers Robots AI* **2020**, vol 7, doi: 10.3389/frobt.2020.00071.
51. WO2025082457 FALL DETECTION AND PREVENTION SYSTEM AND METHOD. Available online: <https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2025082457> (Accessed: Jun. 16, 2025).
52. "US8217795B2 - Method and system for fall detection - Google Patents." Available online: <https://patents.google.com/patent/US8217795B2/en> (Accessed: Jun. 16, 2025).
53. Soaz González, C. "DEVICE, SYSTEM AND METHOD FOR FALL DETECTION," EP 3 796 282 A2, Mar. 21, 2021. Available online: <https://patentimages.storage.googleapis.com/e9/e8/a1/fc9d181803c231/EP3796282A2.pdf#page=19.77> (Accessed: Jun. 16, 2025)
54. Khawandi, S.; Daya, B.; Chauvet, P. Implementation of a monitoring system for fall detection in elderly healthcare. *Proc. Comput. Sci.* **2011**, *3*, 216–220. <https://doi.org/10.1016/j.procs.2010.12.036>.
55. Liao, Y.T.; Huang, C.L.; Hsu, S.C. Slip and fall event detection using Bayesian Belief Network. *Pattern Recognit.* **2012**, *45*, 24–32. <https://doi.org/10.1016/j.patcog.2011.04.017>.
56. Liu, C.L.; Lee, C.H.; Lin, P.M. A fall detection system using k-nearest neighbor classifier. *Expert Syst. Appl.* **2010**, *37*, 7174–7181. <https://doi.org/10.1016/j.eswa.2010.04.014>.
57. Vikman, I.; Nyberg, L.; Korpelainen, R.; Lindblom, J.; Jämsä, T. Comparison of real-life accidental falls in older people with experimental falls in middle-aged test subjects. *Gait Posture* **2012**, *35*, 500–505. <https://doi.org/10.1016/j.gaitpost.2011.11.016>.
58. Zerrouki, N.; Harrou, F.; Sun, Y.; Houacine, A. A data-driven monitoring technique for enhanced fall events detection. *IFAC-PapersOnLine* **2016**, *49*, 333–338. <https://doi.org/10.1016/j.ifacol.2016.07.135>.
59. Martínez-Villaseñor, L.; Ponce, H.; Brieva, J.; Moya-Albor, E.; Núñez-Martínez, J.; Peñafort-Asturiano, C. UP-Fall Detection Dataset: A Multimodal Approach, *Sensors* **2019**, Vol. 19, 1988, doi: 10.3390/S19091988.
60. Charfi, I.; Miteran, J.; Dubois, J.; Atri, M.; Tourki, R. Optimized spatio-temporal descriptors for real-time fall detection: comparison of support vector machine and Ada boost-based classification. *J Electron Imaging* **2013**, vol. 22, no. 4, 041106, doi: 10.1117/1.jei.22.4.041106.
61. Belshaw, M.; Taati, B.; Snoek, J.; Mihailidis, A. Towards a single sensor passive solution for automated fall detection. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society. EMBS, **2011**, 1773–1776. DOI:10.1109/IEMBS.2011.6090506.
62. Charfi, I.; Miteran, J.; Dubois, J.; Atri, M.; Tourki, R. Optimized spatio-temporal descriptors for real-time fall detection: comparison of support vector machine and Adaboost-based classification. *J. Electron. Imaging* **2013**, *22*. <https://doi.org/10.1117/1.JEI.22.4.041106>
63. Fan, Y.; Levine, M.D.; Wen, G.; Qiu, S. A deep neural network for real-time detection of falling humans in naturally occurring scenes. *Neurocomputing* **2017**, *260*, 43–58. <https://doi.org/10.1016/j.neucom.2017.02.082>.
64. Goudelis, G.; Tsatiris, G.; Karpouzis, K.; Kollias, S. Fall detection using history triple features. In: Proceedings of the 8th ACM International Conference on Pervasive Technologies Related to Assistive Environments, 1-3 July, 2015, Corfu Greece, Article No.: 81, Pages 1 – 7. <https://doi.org/10.1145/2769493.2769562>

65. Yu, M.; Yu, Y.; Rhuma, A.; Naqvi, S.M.R.; Wang, L.; Chambers, J.A. 2013. An online one class support vector machine-based person-specific fall detection system for monitoring an elderly individual in a room environment. *IEEE J. Biomed. Heal. Inform.* **2013**, *17*, 1002–1014. <https://doi.org/10.1109/JBHI.2013.2274479>.
66. Debard, G.; Karsmakers, P.; Deschodt, M.; Vlaeyen, E.; Dejaeger, E.; Milisen, K.; Goedemé, T.; Vanrumste, B.; Tuytelaars, T. 2012. Camera-based fall detection on real world data. In: *Outdoor and Large-Scale Real-World Scene Analysis. Lecture Notes in Computer Science*. Dellaert, F., Frahm, J.-M., Pollefeys, M., Leal-Taixé, L., Rosenhahn, B., Eds.; Springer, Berlin, Heidelberg, 2012, 356–375. [https://doi.org/10.1007/978-3-642-34091-8\\_16](https://doi.org/10.1007/978-3-642-34091-8_16).
67. Debard, G.; Mertens, M.; Deschodt, M.; Vlaeyen, E.; Devriendt, E.; Dejaeger, E.; Milisen, K.; Tournoy, J.; Croonenborghs, T.; Goedemé, T.; Tuytelaars, T.; Vanrumste, B. 2016. Camera-based fall detection using real-world versus simulated data: How far are we from the solution? *J. Ambient Intell. Smart Environ* **2016**, *8*, 149–168. <https://doi.org/10.3233/AIS-160369>.
68. Kwolek B.; Kepski, M. Human fall detection on embedded platform using depth maps and wireless accelerometer. *Comput Methods Programs Biomed* **2014**, *117*(3):489-501. doi: 10.1016/j.cmpb.2014.09.005.
69. Vargas, V.; Ramos, P.; Orbe, E.A.; Zapata, M.; Valencia-Aragón, K. Low-Cost Non-Wearable Fall Detection System Implemented on a Single Board Computer for People in Need of Care. *Sensors* **2024**, *Vol. 24*, 5592, doi: 10.3390/S24175592.
70. Li, Y.; Ho, K.C.; Popescu. A microphone array system for automatic fall detection. *IEEE Trans.* **2012**, *Biomed. Eng.* *59*, 1291–1301. <https://doi.org/10.1109/TBME.2012.2186449>.
71. Popescu, M.; Li, Y.; Skubic, M.; Rantz, M. An acoustic fall detector system that uses sound height information to reduce the false alarm rate. In: 30th Annual International IEEE EMBS Conference, 20-23 Aug. 2008, 4628–4631.
72. Salman Khan, M.; Yu, M.; Feng, P.; Wang, L.; Chambers, J. An unsupervised acoustic fall detection system using source separation for sound interference suppression. *Signal Process.* **2015**, *110*, 199–210. <https://doi.org/10.1016/j.sigpro.2014.08.021>.
73. Tao, J.; Turjo, M.; Wong, M.-F.; Wang, M.; Tan, Y.-P. Fall incidents detection for intelligent video surveillance. In: 5th International Conference on Information Communications & Signal Processing, 6-9 Dec, 2005, 1590–1594. <https://doi.org/10.1109/ICICS.2005.1689327>.
74. Vishwakarma, V.; Mandal, C.; Sural. Automatic detection of human fall in video. International Conference on Pattern Recognition and Machine Intelligence, 2007, 616–623. <https://doi.org/10.1007/978-3-540-77046-6>.
75. Wang, S.; Chen, L.; Zhou, Z.; Sun, X.; Dong, J. Human fall detection in surveillance video based on PCANet. *Multimed. Tools Appl.* **2016**, *75*, 11603–11613. <https://doi.org/10.1007/s11042-015-2698-y>.
76. Zhang, Z.; Tong, L.G.; Wang, L. Experiments with computer vision methods for fall detection. In: Proceedings of the 3rd International Conference on Pervasive Technologies Related to Assistive Environments (PETRA '10), 2010, <https://doi.org/10.1145/1839294.1839324>.
77. Zweng, A.; Zambanini, S.; Kampel, M. 2010. *Introducing a statistical behavior model into camera-based fall detection*. In: *Advances in Visual Computing. ISVC 2010. Lecture Notes in Computer Science*. Bebis, G., Boyle, R., Parvin, B., Koracin, D., Chung, R., Hammoud, R., Hussain, M., Kar-Han, T., Crawfis, R., Thalmann, D., Kao, D., Avila, L., Eds.; 2010, Springer, Berlin, Heidelberg, 163–172. [https://doi.org/10.1007/978-3-642-17289-2\\_16](https://doi.org/10.1007/978-3-642-17289-2_16).
78. Senouci, B.; Charfi, I.; Heyrman, B.; Dubois, J.; Miteran, J. Fast prototyping of a SoC-based smart-camera: a real-time fall detection case study. *J. Real-Time Image Process.* **2016**, *12*, 649–662. <https://doi.org/10.1007/s11554-014-0456-4>.
79. De Miguel, K.; Brunete, A.; Hernando, M.; E. Gambao. Home camera-based fall detection system for the elderly. *Sensors* **2017**, *vol. 17*, *no. 12*, doi: 10.3390/s17122864.
80. Hazelhoff, L.; Han, J.; de With, P.H.N. 2008. Video-based fall detection in the home using principal component analysis. In: *Advanced Concepts for Intelligent Vision Systems. ACIVS 2008. Lecture Notes in Computer Science*. Blanc-Talon, J., Bourennane, S., Philips, W., Popescu, D., Scheunders, P., Eds.; Springer, Berlin, Heidelberg, 2008, 298–309. <https://doi.org/10.1007/978-3-540-88458-3>.

81. Feng, W.; Liu, R.; Zhu, M. Fall detection for elderly person care in a vision based home surveillance environment using a monocular camera. *Signal, Image Video Process* **2014**, *8*, 1129–1138. <https://doi.org/10.1007/s11760-014-0645-4>.
82. Foroughi, H.; Aski, B.S.; Pourreza, H. Intelligent video surveillance for monitoring fall detection of elderly in home environments. In: Proceedings of 11th International Conference on Computer and Information Technology, ICCIT 2008, 219–224, <https://doi.org/10.1109/ICCITECHN.2008.4803020>.
83. Horaud, R.; Hansard, M.; Evangelidis, G.; Clément, M. An overview of depth cameras and range scanners based on time-of-flight technologies. *Mach. Vis. Appl.* **2016**, *27*, 1005–1020. <https://doi.org/10.1007/s00138-016-0784-4>.
84. Stone, E.E.; Skubic, M. Fall detection in homes of older adults using the microsoft kinect. *IEEE J. Biomed. Heal. Inform.* **2015**, *19*, 290–301. <https://doi.org/10.1109/JBHI.2014.2312180>.
85. Martínez-Villaseñor, L.; Ponce, H.; Brieva, J.; Moya-Albor, E.; Núñez-Martínez, E.; Peñafort-Asturiano, C. UP-Fall Detection Dataset: A Multimodal Approach, *Sensors* **2019**, *Vol. 19*, no. 9, 1988, doi: 10.3390/S19091988.
86. Toreyin, B.U.; Dedeoglu, Y.; Cetin, A.E. HMM based falling person detection using both audio and video. In: Computer Vision in Human-Computer Interaction. HCI, Lecture Notes in Computer Science. Sebe, N., Lew, M., Huang, T., Eds.; Springer, Berlin, Heidelberg, 2005, 211–220. <https://doi.org/10.1109/SIU.2006.1659753>.
87. Geertsema, E.; Visser, G.H.; Viergever M.A.; Kalitzin S.N. Automated remote fall detection using impact features from video and audio, *Journal of Biomechanics*, **2018**, *88*, 25-32DOI: 10.1016/j.jbiomech.2019.03.007
88. Wu L.; et al., Robust fall detection in video surveillance based on weakly supervised learning. *Neural Networks* **2023**, *vol. 163*, 286–297, doi: 10.1016/J.NEUNET.2023.03.042.
89. Chhetri, S.; Alsadoon, A.; Al-Dala'In, T.; Prasad, P.W.C.; Rashid, T.A.; Maag, A. Deep Learning for Vision-Based Fall Detection System: Enhanced Optical Dynamic Flow. *Comput Intell*, **2021**, *vol. 37*, no. 1, 578–595, doi: 10.1111/coin.12428.
90. Gaya-Morey, F.X.; Manresa-Yee, C.; Buades-Rubio, J.M. Deep learning for computer vision based activity recognition and fall detection of the elderly: a systematic review. *Applied Intelligence* **2024**, *vol. 54*, no. 19, 8982–9007, doi: 10.1007/s10489-024-05645-1.
91. Alhimale, L.; Zedan, H.; Al-Bayatti, A. The implementation of an intelligent and video-based fall detection system using a neural network. *Appl. Soft Comput.* **2014**, *18*, 59–69. <https://doi.org/10.1016/j.asoc.2014.01.024>.
92. Fan, Y.; Levine, M.D.; Wen, G.; Qiu, S. A deep neural network for real-time detection of falling humans in naturally occurring scenes. *Neurocomputing* **2017**, *260*, 43–58. <https://doi.org/10.1016/j.neucom.2017.02.082>.
93. Karpuzov, S.; Kalitzin, S.; Georgieva, O.; Trifonov, A.; Stoyanov T.; Petkov G. Automated remote detection of falls using direct reconstruction of optical flow principal motion parameters. *Sensors* **2025**, *Under Review*; doi: 10.20944/preprints202508.0045.v1
94. Hsieh Y.Z.; Jeng, Y.L. Development of Home Intelligent Fall Detection IoT System Based on Feedback Optical Flow Convolutional Neural Network. *IEEE Access* **2017**, *vol. 6*, 6048–6057. doi: 10.1109/ACCESS.2017.2771389.
95. Huang, C.; Chen, E.; Chung, P. Fall detection using modular neural networks with back-projected optical flow. *Biomed. Eng. Appl. Basis Commun.* **2007**, *19*, 415–424. <https://doi.org/10.1142/S1016237207000525>.
96. Lacuey, N.; Zonjy, B.; Hampson, J.P.; Rani, M.R.S.; Devinsky, O.; Nei, M.; et al. The incidence and significance of periictal apnea in epileptic seizures, *Epilepsia* **2018**, *59*, 573–582, <http://dx.doi.org/10.1111/epi.14006>.
97. Van de Vel, A.; Cuppens, K.; Bonroy, B.; Milosevic, M.; Jansen, K.; Van Huffel, S.; Vanrumste, B.; Lagae, L.; Ceuleman B. Non-EEG seizure-detection systems and potential SUDEP prevention: state of the art, *Seizure* **2013**, *22*, 345–355. <http://dx.doi.org/10.1016/j.seizure.2013.02.012>
98. Baillieul, S.; Revol, B.; Jullian-Desayes, L.; Joyeux-Faure, M.; Tamisier, R.; Pépin, J.-L. Diagnosis and management of sleep apnea syndrome, *Expert Rev. Respir.Med.* **2019**, *13*, 445–557. <http://dx.doi.org/10.1080/17476348.2019.1604226>.

99. Senaratna, CV.; Perret, J.L.; Lodge, C.J.; Lowe, A.J.; Campbell, B.E.; Matheson, M.C.; Hamilton, G.S.; Dharmage, S.C. Prevalence of obstructive sleep apnea in the general population: a systematic review, *Sleep Med. Rev.* **2017**, *34*, 70–81. <http://dx.doi.org/10.1016/j.smrv.2016.07.002>.
100. Zaffaroni, A.; Kent, B.; O'Hare, E.; Heneghan, C.; Boyle, P.; O'Connell, G.; Pallin, P. M.; De Chazal, W.T. McNicholas, Assessment of sleep-disordered breathing using a non-contact bio-motion sensor, *J. Sleep Res.* **2013**, *22*, 231–236. <http://dx.doi.org/10.1111/j.1365-2869.2012.01056.x>.
101. Castro, I.D.; Varon, C.; Torfs, T.; van Huffel, S.; Puers, R.; van Hoof, C. Evaluation of a multichannel non-contact ECG system and signal quality algorithms for sleep apnea detection and monitoring, *Sensors* **2018**, *18* 1–20. <http://dx.doi.org/10.3390/s18020577>.
102. Hers, V.; Corbugy, D.; Joslet, I.; Hermant, P.; Demarteau, J.; Delhougne, B.; Vandermoten, G.; Hermanne, J.P. New concept using Passive Infrared (PIR) technology for a contactless detection of breathing movement: a pilot study involving a cohort of 169 adult patients. *J. Clin. Monit. Comput.* **2013**, *27*, 521–529. <http://dx.doi.org/10.1007/s10877-013-9457-2>.
103. Garn, H. Kohn, B.; Wiesmeyr, C.; Dittrich, K.; Wimmer, M.; Mandl, M.; Kloesch, G.; Boeck, M.; Stefanic, A.; Seidel, S. 3D detection of the central sleep apnoea syndrome, *Curr. Dir. Biomed. Eng.* **2017**, *3*, 829–833. <http://dx.doi.org/10.1515/cdbme-2017-0174>.
104. Nandakumar, R.; Gollakota, S.; Watson, N. Contactless sleep apnea detection on smart phones, Proc. 13th Annu. Int. Conf. Mob. Syst. Appl. Serv. - MobiSys' 15, Florence Italy, May 18-22, 2015, 45–57, <http://dx.doi.org/10.1145/2742647.2742674>.
105. Al-Naji, A.; Gibson, K.; Lee, S.-H.; Chahl, J. Real time apnoea monitoring of children using the microsoft kinect sensor pilot study. *Sensors* **2017**, *17*, 286. <http://dx.doi.org/10.3390/s17020286>.
106. Yang, C.; Cheung, G.; Stankovic, V.; Chan, K.; Ono, N. Sleep apnea detection via depth video and audio feature learning, *IEEE Trans. Multimed.* **2017**, *19*, 822–835. <http://dx.doi.org/10.1109/TMM.2016.2626969>.
107. Wang, C.W. Hunter, A.; Gravill, N.; Matusiewicz, S. Unconstrained video monitoring of breathing behavior and application to diagnosis of sleep apnea, *IEEE Trans. Biomed. Eng.* **2014**, *61*, 396–404. <http://dx.doi.org/10.1109/TBME.2013.2280132>.
108. Sharma, S.; Bhattacharyya, S.; Mukherjee, J.; Purkait, P.K.; Biswas, A.; Deb, A.K. Automated detection of newborn sleep apnea using video monitoring system, proceedings: Eighth Int. Conf. Adv. Pattern Recognit. 4-7 January, 2015, England, 1–6. <http://dx.doi.org/10.1109/ICAPR.2015.7050675>.
109. Jorge, J.; Villarreal, M.; Chaichulee, S.; Guazzi, A.; Davis, S.; Green, G.; McCormick, K.; Tarassenko, L. Non-contact monitoring of respiration in the neonatal intensive care unit, Proc. - 12th IEEE Int. Conf. Autom. Face Gesture Recognit., 2017, 286–293, <http://dx.doi.org/10.1109/FG.2017.44>.
110. Cattani, L.; Alinovi, D.; Ferrari, G.; Raheli, R.; Pavlidis, E.; Spagnoli, C.; Pisani, F.; Monitoring F. infants by automatic video processing: a unified approach to motion analysis, *Comput. Biol. Med.* **2017**, *80*, 158–165, <http://dx.doi.org/10.1016/j.compbiomed.2016.11.010>.
111. Koolen, N.; Decroupet, O.; Derymaeker, A.; Jansen, K.; Vervisch, J.; Matic, V.; Vanrumste, B.; Nauelaers, G.; Van Huffel, S.; De Vos, M. Automated respiration detection from neonatal video data, *Proc. 4th Int. Conf. Pattern Recognit. Appl. Methods* **2015**, vol. 2, 164–169. <http://dx.doi.org/10.5220/0005187901640169>.
112. Li, M.H.; Yadollahi, A.; Taati, B. Noncontact vision-based cardiopulmonary monitoring in different sleeping positions. *IEEE J. Biomed. Heal. Inf.* **2017**, *21*, 1367–1375, <http://dx.doi.org/10.1109/JBHI.2016.2567298>.
113. Horn, B.K.P.; Schunck, B.G. Determining optical flow, *Artif. Intell.* **1981**, *17*, 185–203, [http://dx.doi.org/10.1016/0004-3702\(93\)90173-9](http://dx.doi.org/10.1016/0004-3702(93)90173-9).
114. Groote, A.; Wantier, M.; Cheron, G.; Estenne, M.; Paiva, M. Chest wall motion during tidal breathing, *J. Appl. Physiol.* **1997**, *83*, 1531–1537, <http://dx.doi.org/10.1152/jappl.1997.83.5.1531>.
115. Geertsema, E.E.; Visser, G.H.; Sander, J.W.; Kalitzin, S.N. Automated non-contact detection of central apneas using video. *Biomedical Signal Processing and Control* **2020**, *55*, 101658. <https://doi.org/10.1016/j.bspc.2019.101658>
116. Petkov, G.; Mladenov, N.; Kalitzin, S. Integral scene reconstruction from general over-complete sets of measurements with application to explosions localization and charge estimation. *Int. Journal of Computer Aided Engineering*, **2013**, *20*, 95-110

117. Higham J.E.; Isaac, O.S.; Rigby, S.E. Optical flow tracking velocimetry of near-field explosion *Meas. Sci. Technol.* **2022**, *33* 047001 DOI 10.1088/1361-6501/ac4599
118. Yilmaz, A.; Javed, O.; Shah, M. Object tracking: A survey. *ACM Comput. Surv.* **2006**, *38*, 13-es.
119. Li, X.; Hu, W.; Shen, C.; Zhang, Z.; Dick, A.; Hengel, A.V.D. A survey of appearance models in visual object tracking. *ACM Trans. Intell. Syst. Technol.* **2013**, *4*, 1–48.
120. Chen, F.; Wang, X.; Zhao, Y.; Lv, S.; Niu, X. Visual object tracking: A survey. *Comput. Vis. Image Underst.* **2022**, *222*, 103508.
121. Ondrašovič, M.; Tarábek, P. Siamese visual object tracking: A survey. *IEEE Access* **2021**, *9*, 110149–110172.
122. Farag, W.; Saleh, Z. An advanced vehicle detection and tracking scheme for self-driving cars. In Proceedings of the 2nd Smart Cities Symposium (SCS 2019), Bahrain, 24–26 March 2019; IET: Stevenage, UK, 2019.
123. Gupta, A.; Anpalagan, A.; Guan, L.; Khwaja, A.S. Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues. *Array* **2021**, *10*, 100057.
124. Lipton, A.J.; Reartwell, C.; Haering, N.; Madden, D. Automated video protection, monitoring & detection. *IEEE Aerosp. Electron. Syst. Mag.* **2003**, *18*, 3–18.
125. Wang, W.; Gee, T.; Price, J.; Qi, H. Real time multi-vehicle tracking and counting at intersections from a fisheye camera. In Proceedings of the 2015 IEEE Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 5–9 January 2015; IEEE: Piscataway, NJ, USA, 2015.
126. Kim, H. Multiple vehicle tracking and classification system with a convolutional neural network. *J. Ambient Intell. Humaniz. Comput.* **2022**, *13*, 1603–1614.
127. Deori, B.; Thounaojam, D.M. A survey on moving object tracking in video. *Int. J. Inf. Theory* **2014**, *3*, 31–46.
128. Mangawati, A.; Leesam, M.; Aradhya, H.R. Object Tracking Algorithms for video surveillance applications. In Proceedings of the 2018 International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur, India, 3–5 April 2018; IEEE: Piscataway, NJ, USA, 2018.
129. Gilbert, A.; Bowden, R. Incremental, scalable tracking of objects inter camera. *Comput. Vis. Image Underst.* **2008**, *111*, 43–58.
130. Yeo, H.-S.; Lee, B.-G.; Lim, H. Hand tracking and gesture recognition system for human-computer interaction using low-cost hardware. *Multimed. Tools Appl.* **2015**, *74*, 2687–2715.
131. Fagiani, C.; Betke, M.; Gips, J. Evaluation of Tracking Methods for Human-Computer Interaction. In Proceedings of the WACV, Orlando, FL, USA, 3–4 December 2002.
132. Hunke, M.; Waibel, A. Face locating and tracking for human-computer interaction. In Proceedings of the 1994 28th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 31 October–2 November 1994; IEEE: Piscataway, NJ, USA, 1994.
133. Karpuzov S.; Petkov, G.; Ilieva, S.; Petkov A.; Kalitzin S. Object Tracking Based on Optical Flow Reconstruction of Motion-Group Parameters. *Information* **2024**, *15*, 296. <https://doi.org/10.3390/info15060296>
134. Doyle, D.D.; Jennings, A.L.; Black, J.T. Optical flow background estimation for real-time pan/tilt camera object tracking. *Measurement* **2014**, *48*, 195–207.
135. Hussein, S. A Survey of Optical Flow Techniques for Object Tracking. Bachelor's Thesis, Tampere University, Tampere, Finland, 2017.
136. Zhang, P.; Tao, Z.; Yang, W.; Chen, M.; Ding, S.; Liu, X.; Yang, R.; Zhang, H. Unveiling personnel movement in a larger indoor area with a non-overlapping multi-camera system. *arXiv* **2021**, *arXiv:2104.04662*.
137. Porikli, F.; Divakaran, A. Multi-camera calibration, object tracking and query generation. In Proceedings of the 2003 International Conference on Multimedia and Expo. ICME'03, Baltimore, MD, USA, 6–9 July 2003; Proceedings (Cat. No. 03TH8698); IEEE: Piscataway, NJ, USA, 2003.
138. Dick, A.R.; Brooks, M.J. A stochastic approach to tracking objects across multiple cameras. In Proceedings of the Australasian Joint Conference on Artificial Intelligence, Cairns, Australia, 4–6 December 2004; Springer: Berlin/Heidelberg, Germany, 2004.
139. Yang, F.; Odashima, S.; Masui, S.; Kusajima, I.; Yamao, S.; Jiang, S. Enhancing Multi-Camera Gymnast Tracking Through Domain Knowledge Integration. *IEEE Trans. Circuits Syst. Video Technol.* **2024**, *34*, 13386–13400.

140. Amosa, T.I.; Sebastian, P.; Izhar, L.I.; Ibrahim, O.; Ayinla, L.S.; Bahashwan, A.A.; Bala, A.; Samaila, Y.A. Multi-camera multi-object tracking: A review of current trends and future advances. *Neurocomputing* **2023**, *552*, 126558.
141. Cherian, R.; Jothimani, K.; Reeja, S. A Review on Object Tracking Across Real-World Multi Camera Environment. *Int. J. Comput. Appl.* **2021**, *174*, 32–37.
142. Yang, F.; Odashima, S.; Yamao, S.; Fujimoto, H.; Masui, S.; Jiang, S. A unified multi-view multi-person tracking framework. *Comput. Vis. Media* **2024**, *10*, 137–160.
143. Karpuzov, S.; Petkov, G.; Kalitzin, S. Multiple-Camera Patient Tracking Method Based on Motion-Group Parameter Reconstruction. *Information* **2025**, *16*, 4. <https://doi.org/10.3390/info16010004>
144. Fei, L.; Han, B. Multi-object multi-camera tracking based on deep learning for intelligent transportation: A review. *Sensors* **2023**, *23*, 3852.
145. Elmenreich, W. An introduction to sensor fusion. *Vienna Univ. Technol. Austria* **2002**, *502*, 1–28.
146. Fung, M.L.; Chen, M.Z.; Chen, Y.H. Sensor fusion: A review of methods and applications. In Proceedings of the 2017 29th Chinese Control and Decision Conference (CCDC), Chongqing, China, 28–30 May 2017.
147. Yeong, D.J.; Velasco-Hernandez, G.; Barry, J.; Walsh, J. Sensor and sensor fusion technology in autonomous vehicles: A review. *Sensors* **2021**, *21*, 2140. [CrossRef]
148. Yu, L.; Ramamoorthi, R. Learning Video Stabilization Using Optical Flow, In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle WA, June 13–19, 2020, 8159–8167,
149. Chang, I.-Y.; Hu, W.-F.; Cheng, M.-H.; Chang, B.-S. Digital image translational and rotational motion stabilization using optical flow technique. *IEEE Transactions on Consumer Electronics*, **2002**, *vol. 48*, *no. 1*, 108–115, Feb. 2002, doi: 10.1109/TCE.2002.1010098.
150. Deng, D.; Yang, D.; Zhang, X.; Dong, Y.; Liu, C.; Shen, Q. Real-Time Image Stabilization Method Based on Optical Flow and Binary Point Feature Matching, *Electronics* **2020**, *9*, 198; doi:10.3390/electronics9010198

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.