On the Vector Space in Photoplethysmography Imaging

Christian S. Pilz
CanControls, Aachen, Germany

Ibtissem Ben Makhlouf
CanControls, Aachen, Germany

Vladimir Blazek
RWTH Aachen University, Germany

Steffen Leonhardt
RWTH Aachen University, Germany

Abstract

We study the vector space of visible wavelength intensities from face videos widely used as input features in Photoplethysmography Imaging (PPGI). Based upon theoretical principles of group invariance in the Euclidean space, we derive a change of the topology where the corresponding distance between successive measurements is defined as geodesic on a Riemannian manifold. This lower dimensional embedding of the sensor signal unifies the invariance properties with respect to translation of the features as discussed by several former approaches. The resulting operator acts implicitly on the feature space without requiring any kind of prior knowledge and parameter tuning. The resulting features time varying quasi-periodic shaping naturally occurs in form of the canonical state space representation according to the known diffusion process of blood volume changes. This reduces drastically computational complexity and consequently simplifies the implementation. Experiments from face videos on two public databases have shown a competitive estimation performances of heart rate and robustness in comparison with already available methods.

1. Introduction

About 70 years ago, short after the end of second World War, Norbert Wiener published his sociopolitical often controversial discussed work *Cybernetics or Control and Communication in the Animal and the Machine* [40]. Much of the technology we take for granted today was developed during this period and is till now contributing at shaping our future. Worthwhile emphasizing is the impact of fast emerging progress in the computational interpretation of sensors, signals and systems of our daily lives. The significant progress observed in many areas gives rise to the question of how far can artificial intelligence reach at least the same performance.

As part of natural social interaction the human face with its contingent of verbal activity, non-verbal behavior as well as its appearance is the main source for the majority of interpretable signals for computer systems. In contrast to the non-verbal behavioral signals the facial appearance is able to provide changes in peripheral nervous system as well as central nervous system surrogate states by the analysis of skin blood perfusion. Because of minor intensity changes, in general not perceivable by human eyes, opto-electronic has be used to access this kind of information. Under the context of non-obtrusive remote measurability this new technology enables acquisitions of biological and cognitive human data under exceptional situations potentially leading to new findings and application fields. Basically, nearly every task targeting on the specific dependent variables, formally captured by cable mounted sensors sticked onto human skin, can be sensed by the camera-based counterpart as well. The possible range of new applications and analysis legitimates to name its potential as ground breaking. The topic is traditionally anchored in the medical sciences. However, currently the focus elementary changed to direction of computer vision. Here the role of physiological states has a large impact. It is primarily used during human state computing tasks, where the radiation unnoticeable transports information from face without contact holding states of affective nature.

During the last years measuring blood volume changes and heart rate measurements from facial images gained attention and was the topic of range of top computer vision conferences [17, 26, 35, 27, 28, 9] frequently. Most of these contributions focus on how to cope with motion like head pose variations and facial expressions since any kind of motion on a specific skin region of interest will destroy the raw signal in a way that no reliable information can be extracted anymore. Beside being able to estimate vitality parameters like heart rate and respiration, the functional

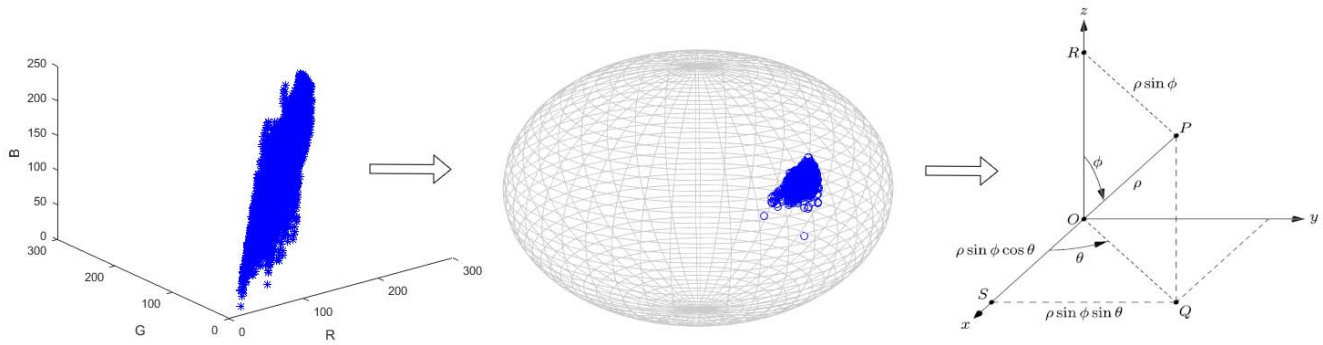


Figure 1: The PPGI Signal as distribution of pixel intensities in the Euclidean RGB space is embedded in a lower dimensional space on the unit sphere. On the left, the skin pixel intensities are shown in the RGB coordinate system. In the middle, the pixel intensities are mapped onto the unit sphere. On the right, their schematic representation as spherical coordinates is shown.

survey of wounds as well as quantification of allergic skin reaction [4] are further topics of discovered features of skin blood perfusion analysis. Recently, prediction of emotional states, stress [22, 30, 6], fatigue [33] and sickness [13] aroused great interest in this area, pushing the focus of this technology further towards human-machine interaction.

In contrast to the genuine medical use-case of the technology, in computer vision and human-machine interaction we can not expect any cooperative behavior of the user without introducing lack of convenience and a reduction of the general user acceptance. Further, beyond any well-tempered clinical and laboratory like scenarios, the majority application will face frequently strong challenging environmental changes and differences. This arises an urgent demand on computing robust models and features able to preserve intended information under disturbances and uncertainties. To reach such a goal, an in-depth understanding of these physical properties is fundamentally required.

The main contribution of this work is a mathematical analysis of the PPGI's feature space determined over facial skin pixels. The general aim is to study the properties and behavior of the features with respect to the influence of the actions of the group acting on this space when induced by natural head and face motion. As result, a new feature operator is developed and evaluated against common operators on various data sets. These sets are collected to explicitly study the influence of typical nuisance attributes. To support an efficient dissemination and to speed up the research progress in this field, we have encapsulated all feature operators and the reference implementation of the stochastic model of blood volume changes into a new MATLAB toolbox. The code is already public available via the author's personal web page. The toolbox provides the code to reproduce all results presented in this work

This work is organized as follows. From the historical genuine up to the development of the state-of-the-art in computer vision and bio-medical engineering, the methodology of heart rate estimation from face videos will be reviewed. Followed by theoretical aspects, the feature space will be analyzed and the proposed methodology mathematically described. In Figure 1 the main steps of the underlying signal processing is illustrated. Based upon an extensive evaluation on two different databases the results will be presented and finally discussed.

2. Related Work

The earlier history of Photoplethysmography, short PPG, dates back to the late first half of the 20th century, when the two scientists Molitor and Kniazak [25] recorded peripheral circulatory changes in animals. A year later, Hertzman [14] introduced the term Photoelectric Plethysmograph for "the amplitude of volume pulse as a measure of the blood supply of the skin". Hertzman's instrumentation consisted in a tungsten arc lamp and a photomultiplier tube. Meanwhile a similar experiment has been presented in a medical journal by Mathes [20]. Although there is no precise information about the first origin of this experiment, the main contribution is in the literature accredited to Hertzman.

Around 50 years later the advancement to the classical PPG, the camera based PPGI (with I for Imaging) method, was introduced by the pioneering work of Blazek [5]. The basic principle behind the measurement of blood volume changes in the skin by means of PPGI (as well as PGG too) is the fact that the oxygen binding ferrous protein complex hemoglobin in the blood absorbs specific

frequency bands of light many times stronger than the remaining skin tissues. Accordingly, tiny intensity changes can be observed over specific frequency bands (e.g. the density of spectral lines of the emission spectra of iron) as oscillation caused by the quasi periodic heart rhythm. In PGG a part of the skin surface is illuminated by dedicated light sources like illumination panels consisting of LEDs. In PPGI a common CCD camera is used as detector and the illumination can be ordinary common ambient light for which the intensity of backscattered optical radiation, eq. reflected light, is calculated [2, 15, 36].

In general the computational pipeline to determine vitality parameters and their derivatives from blood volume changes can be regarded as a classical signal processing chain. Typically, from a skin region of interest (ROI) features are calculated, filtered and analyzed by spectral methods [15, 29, 36]. The first published visualization of pulsatile skin perfusion patterns in the time and frequency domain is given by Blazek [5]. However motion of the skin ROI [15] and micro motion of the head due to cardiac activity [3, 23] inherently induces artifacts into the extracted signal, especially when lighting is neither uniform nor orthogonal [24]. Canceling motion artifacts during signal processing became one of the most important aspect for reliable skin blood perfusion measurements. An early idea of skin ROI motion compensating is to track every skin pixel position by optical flow methods directly in the image plane [15]. However this doesn't account for any change of illumination. Poh *et al.* [29] proposed to extract motion components in the signal by blind source separation using Independent Component Analysis (ICA) over the different color channels. Wedekind *et al.* [39] compared ICA in multiple setting and Principal Component Analysis and showed limitations of these transformations. Further, the ICs cannot be obtained in a deterministic order [8]. A solution to this problem is discussed by Macwan *et al.* [18]. Tarassenko *et al.* [34] tried to cope with light flicker by using an auto-regressive modeling and pole cancellation. De Haan and Jeanne [10] and De Haan and Van Leest [11] proposed to map the PPGI-signals by linear combination of RGB data to a direction that is orthogonal to motion induced artifacts. An alternative approach, which does not require skin-tone or pulse-related priors in contrast to the channel mapping algorithms, determines the spatial subspace of skin-pixels and measures its temporal rotation for signal extraction [38]. Tulyakov *et al.* [35] proposed matrix completion to jointly estimate reliable regions and heart rate estimates whereby Li *et al.* [17] applied an adaptive least square approach to extract robust pulse frequencies. Both reported performance gains similar to De Haan and Jeanne [10]. Interestingly, they used the often criticized compressed videos of the MAHNOB-HCI database [32] during their experiments. This leaves reasonable doubts on the validity of results, since it is well known that any kind of image compression will destroy the underlying tiny perfusion signal [21]. Wang *et al.* [37] reported an orthogonal behavior of skin color and motion artifacts derived by optical properties but introduced a static projection operator for feature transformation. They represented their results on private data. An entirely different model was introduced by Pilz *et al.* [27]. Here, the quasi-periodic nature of the blood volume changes is modeled as stochastic resonator based on a diffusion process. A group theoretic deviated feature transformation for motion compensation was furthermore introduced by Pilz *et al.* [28]. Chen and McDuff [9] made use of Deep Learning methods to outperform recent algorithms using a convolutional network architecture (CNN) for modelling motion representation. However, they also reported some results on the compressed videos of the MAHNOB-HCI database and they didn't provide their CNN implementation or at least the trained model.

3. Methodology

Subject motion and fast strong changes of illumination alter the distribution of pixel intensity negatively making it quite difficult to extract skin perfusion signals from video images. However, it is assumed that the perfusion signals exits in both cases and and that those signals are somehow in relation with the distribution of intensity belonging to motion forces. We use the basic principles of the Hilbert projection theorem as the foundation of least-squares approximation to find a mathematical model for this relation. This theorem states that if $\mathcal{V}$ is a closed subspace of the Hilbert Space $\mathcal{H}$ and $x \in \mathcal{H}$, then there is a unique element $\hat{x} \in \mathcal{V}$ such that

$$\|x - \hat{x}\| = \inf_{y \in \mathcal{V}} \|x - y\| \tag{1}$$

and only if $\hat{x} \in \mathcal{V}$ and $(x - \hat{x}) \in \mathcal{V}^{\perp}$ where $\mathcal{V}^{\perp}$ is the orthogonal complement to $\mathcal{V}$ in $\mathcal{H}$. Then, the current paradigm in understanding PPGI signal components in the feature space assumes that blood volume changes exist in a lower dimensional space where this space is orthogonal to any kind of motion induces signal components. Derived from optical properties described by De Haan *et al.* and Wang *et al.* [10, 37] and using group theoretic principals introduced by Pilz *et al.*, [28] this can be expressed as $\vec{c} = \vec{p} + \vec{m}$. $\vec{p}$ and $\vec{m}$ are two orthogonal statistically linear independent vectors with $\vec{p} \cdot \vec{m}^T = 0$. This principal is illustrated in Figure 2. In the following, we first recall the group theoretic principals behind motion robust sensing of blood volume changes introduced by Pilz *et al.* [28]. Based on the analysis of the properties of the resulting linear operator we demonstrate that there exists an equivalent implicit operator. This op-

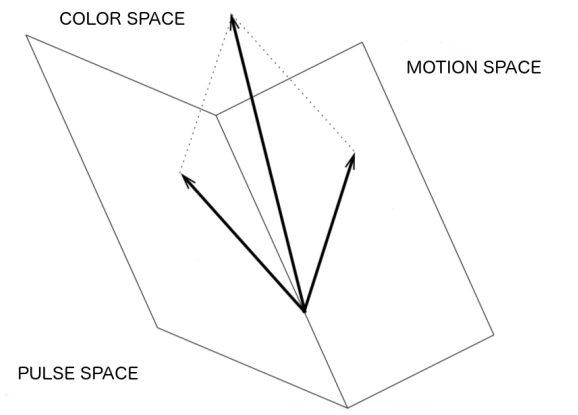


Figure 2: The actual state-of-the-art paradigm in understanding major PPGI signal components in the feature space. With $\vec{c} = \vec{p} + \vec{m}$, $\vec{p}$ and $\vec{m}$ are two orthogonal vectors with $\vec{p} \cdot \vec{m}^T = 0$.

erator maps the observation onto an embedded Riemannian submanifold of the Euclidean space. In addition we show that the corresponding directional statistics evolve in form of the previously published diffusion process model as function of time [27].

3.1. Basic Principals of Group Invariance

Consider a finite topological group

$$\mathcal{G} = \{G_1, ..., G_M\} \tag{2}$$

of M distinct actions on a topological space

$$\mathbb{X}, G_i : \mathcal{X} \to \mathcal{X}. \tag{3}$$

A real valued function $f(x)$ on $\mathcal{X}$ is said to be invariant under $\mathcal{G}$ if

$$f(Gx) = f(x) \text{ for } G \in \mathcal{G} \tag{4}$$

Regarding a common optical sensor signal $\{\vec{p_i} : i = 1, ..., m\}$

$$\vec{p} \in \mathbb{R}^n = \{Red, Green, Blue\}, n = 3 \tag{5}$$

as spatial expectation over a skin operator s and function of time t

$$\vec{x}(t) = \int_0^\infty \mathbb{E}[\{\vec{p_i} \mid s(\vec{p_i})\}]\mathrm{d}t. \tag{6}$$

We assume this multivariate observation is given by a normal distribution

$$\vec{x}(t) \sim \mathcal{N}(\vec{\mu}, \vec{\sigma}^2). \tag{7}$$

Local invariance of blood volume changes as infinitesimal generators of the special Euclidean group SE(3) for each input feature $\vec{x}(t)$ under transformations of a differentiable local group of local transformations $\mathcal{L}_T$ [31]

$$\frac{\partial}{\partial T}\Big|_{T=0} = f(\mathcal{L}_T, \vec{x}(t)) = 0 \tag{8}$$

can be approximately enforced by minimizing the regularizer

$$\frac{1}{l}\sum_{j=1}^{l}(\frac{\partial}{\partial T}\Big|_{T=0} f(\mathcal{L}_T, x_j))^2. \tag{9}$$

For the covariance matrix of the observation $X := \{x_i : i = 1, ..., l\}$ with respect to the transformations $\mathcal{L}_T$

$$C := \frac{1}{l}\sum_{j=1}^{l}(\frac{\partial}{\partial T}\Big|_{T=0}\mathcal{L}_T, x_j)(\frac{\partial}{\partial T}\Big|_{T=0}\mathcal{L}_T, x_j)^\top \tag{10}$$

and the corresponding symmetric eigenvalue problem

$$CV = V\Lambda \tag{11}$$

we find an operator P with corank $k = 1$ with

$$\lim_{l\to\infty} P = I - VV^\top \tag{12}$$

and the corresponding feature vector

$$\tilde{x} = P \cdot x. \tag{13}$$

The observation $\tilde{X} := \{\tilde{x_i} : i = 1, ..., l\}$ defines the null space of the projection operator P

$$H_P = N(P) \tag{14}$$

In general, we are particulary interested in the properties of the projection operator P and the resulting linear subspace H_P since the direction of V is assumed to carry most of the PPGIs motion signal component and the subspace H_P the desired quasi periodic perfusion component. The projection P removes the direction of the largest variance and emphasizes the less varying directions. It should be clear that the computation of the projection is an explicit operator which has to be estimated using the observation X. Inherently, this produces a time lag of l samples into the past. Therefore, a better choice of incorporating invariance to the feature space would be to define an implicit operator.

3.2. The embedded Riemannian submanifold

It is a known from the theory of Banach algebras that the spectral radius ρ of any $A \in \mathcal{M}_n$ is given by Gelfand's formula [12]

$$\rho(A) = \lim_{n\to\infty} \|A^n\|^{1/n} \tag{15}$$

for any matrix norm $\|\cdot\|$ on $\mathcal{M}_n$. For a matrix, the spectrum is just the collection of eigenvalues with $\rho(A) := max\{|\lambda|, \lambda \text{ eigenvalue of } A\}$. For the projection operator P, where V is represented by the eigenvector with the largest eigenvalue, this implies a reduction of the spectral radius of the observation $X := \{x_i : i = 1, ..., l\}$

$$\rho(\tilde{X}) < \rho(X). \tag{16}$$

Let λ be an eigenvalue of X, and let $v \neq 0$ be a corresponding eigenvector. From $Xv = \lambda v$, we have $XV = \lambda V$ where $V := [v \mid ... \mid v] \in \mathcal{M}_n \setminus \{0\}$. It follows $|\lambda| \|V\| = \|\lambda V\| = \|XV\| \leq \|X\|\|V\|$, and simplifying by $\|V\| (> 0)$ gives $|\lambda| \leq \|X\|$. Taking the maximum over all eigenvalues λ results in $\rho(X) \leq \|X\|$.

Now, regarding the genuine optical sensor signal $\vec{p}_i$ relative to the spectral radius of its vector elements $\rho(\vec{p_{ij}})$ with

$$\vec{x}_{\mathcal{M}}(t) = \int_0^{\infty} \mathbb{E}[\{\frac{\vec{p}_{ij}}{\|\vec{p_{ij}}\|} \mid s(\vec{p_i})\}]\mathrm{d}t \tag{17}$$

and $X_{\mathcal{M}} := \vec{x}_{\mathcal{M}}(t)$ results in

$$\rho(X_{\mathcal{M}}) < \rho(X) \Rightarrow \rho(X_{\mathcal{M}}) \equiv \rho(\tilde{X}) \tag{18}$$

with a distribution on $\mathbb{S}^2 = \{\vec{x}_{\mathcal{M}} \in \mathbb{R}^3 \mid \|\vec{x}_{\mathcal{M}}\| = 1\}$, the unit sphere as embedded Riemannian submanifold of the Euclidean space $\mathbb{R}^3$.

Intuitively, for real valued observations $\{x_i : i = 1, ..., m \in \mathbb{R}\}$, the mean is given by $\mu = \frac{1}{m}\sum_{i=1}^{m} x_i$. However, for the more general settings $\{x_i : i = 1, ..., n \in \mathcal{M}\}$ we do not have the possibility to find such a closed form solution. The solution has to be solved for the following optimization problem

$$\mu \in \arg\min_{x \in \mathcal{M}} \frac{1}{m} \sum_{i=1}^{m} d^2_{\mathcal{M}}(x, x_i) = \arg\min_{x \in \mathcal{M}} F(x). \tag{19}$$

by using a gradient descent algorithm. The resulting μ is called the Riemannian center of mass or Karcher mean [16]. The optimality conditions is give by

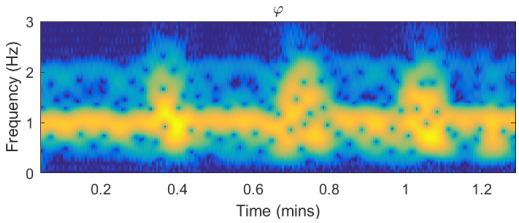
$$0 \overset{!}{=} \frac{2}{m} \sum_{i=1}^{n} log_{x^*} x_i = \nabla F(x^*) \in T_{x^*}\mathcal{M}. \tag{20}$$

With the involved gradient derived by Asfari [1], we are able to solve the objective by performing the descent steps from some initial point $x^{(0)}$ given by

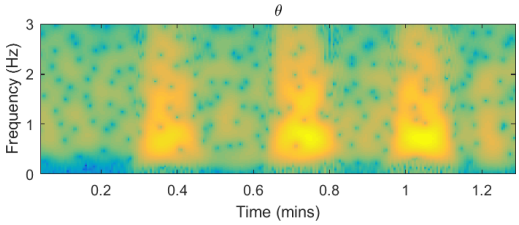
$$x^{(k+1)} = exp_{x^{(k)}}\left(-s_k \nabla_{\mathcal{M}} F\left(x^k\right)\right). \tag{21}$$

The directional statistics are represented by the von Mises-Fisher distribution with the probability density given by

$$p(\vec{x}_{\mathcal{M}}; \vec{\mu}, \kappa) := c_p(\kappa)\, e^{\kappa \vec{\mu}^T \vec{x}_{\mathcal{M}}} \tag{22}$$



(a) The coordinate φ shows a strong pulse frequency and a small portion of motion energy.



(b) The coordinate θ shows sensor noise and the major portion of motion energy.

Figure 3: The spectrum of the spherical coordinates φ and θ computed over a head rotation sequence.

with the normalization constant given by

$$c_p(\kappa) = \frac{\kappa^{p/2-1}}{(2\pi)^{p/2} I_{p/2-1}(\kappa)} \tag{23}$$

where $I_s(\kappa)$ denotes the modified Bessel function of the first kind [19].

Regarding the embedded dimensions of the unit sphere given by its spherical coordinates with

$$r = \sqrt{\vec{x}^2_{\mathcal{M}_1} + \vec{x}^2_{\mathcal{M}_2} + \vec{x}^2_{\mathcal{M}_3}} = \|\vec{x}_{\mathcal{M}}\| = 1, \tag{24}$$

$$\theta = arccos\frac{\vec{x}_{\mathcal{M}_3}}{r} = arccos\vec{x}_{\mathcal{M}_3}, \tag{25}$$

and

$$\varphi = arctan\frac{\vec{x}_{\mathcal{M}_2}}{\vec{x}_{\mathcal{M}_1}}. \tag{26}$$

The spherical coordinate φ evolves according to the principals of the stochastic differential equation given by

$$\frac{d^2 c_n(t)}{dt^2} = -(2\pi n f(t))^2 c_n(t) + e_n(t) \tag{27}$$

whereby θ can be expressed as part of a Wiener process given by

$$\frac{\mathrm{d}^2\theta(t)}{\mathrm{d}t^2} = w(t). \tag{28}$$

(a) UBFC-RPPG

(b) LGI Multi-Session

Figure 4: Example images taken from the UBFC-RPPG and LGI Multi-Session database.

For the details on the derivation of the diffusion process we refer to the genuine work of Pilz *et al.* [27]. In Figure 3a and 3b the spectrum of the spherical coordinates φ and θ computed over a head rotation sequence taken from the LGI database [28] is shown.

4. Experiments

The evaluation of the described feature operator was conducted by comperative experiments using different available methods on two public available databases. We decided to compare the operator against the baseline green channel expectation [15, 36], the Spatial Subspace Rotation (SSR) [38], the Projection Orthogonal to Skin (POS) [37] and the Local Group Invariance (LGI) [28] method. All experiments were executed on the French UBFC-RPPG [7] and the German LGI Multi-Session face video database [28]. In Figures 4a and 4b some representative images of these two databases are illustrated. The UBFC and the LGI database were created using a custom C++ application for video acquisition with a simple low cost Logitech webcam and a CMS50E transmissive pulse oximeter to obtain the ground truth PPG data comprising the PPG waveform. The total amount of face video recordings yields to 150 sequences containing several couple of minutes respectively. During the recordings, the subjects of the UBFC trials performed moderate face and head motions under indoor environments whereby the LGI recordings contain scenarios from resting and head rotation over sport activities to natural outdoor conversations. Therefore, the evaluation concept ranges from cooperative to challenging scenarios which should be become noticeable in form of the prediction accuracy results.

The primary signal processing procedure is selected to be equal for every approach and database. For each video frame a common Viola-Jones face finder was used to preselect the region of interest. A simple skin operator was applied onto the region by thresholding the blue- and red-difference chroma components. For the set of obtained RGB-pixels the different approach specific operators were computed and stored as time series for further spectral processing and interpretation. Each signal obtained by the different algorithms was band-filtered in the range between 0.5 and 2.5 Hz. All filtered signals were then analyzed by standard Fourier based spectral method with windows size of 256 samples and overlap of 90 percent. A maximum peak energy criterion was applied over the spectral traces to determine the heart rate. All PPG ground truth signals were analyzed in the same way but initially resampled to the camera frame rate. Correlation coefficients were computed against the PPG reference heart rate together with the root-mean-square error (RMSE) for each user and algorithm respectively. We did not perform a Signal-to-noise ratio (SNR) comparison as proposed by De Haan and Jeanne [10] and often used by several other authors [11, 37, 9]. This might be useful on short video sequences with a more or less stationary frequency behavior of blood volume changes. A prospective consideration of a SNR metric for system evaluation should at least include information about it's variance computed on short term spectra.

Table 1 presents the overview of heart rate prediction accuracy of the different operators on the different data sets. Compared to the baseline green channel, SSR and POS approach, the LGI and the proposed spherical operator (SPH) are more robust especially under motion scenarios. Although the SPH operator cannot outperform the LGI approach in most cases, it's accuracy is operating in a very similar range. However, in case of fully uncontrolled scenarios with changing illumination as well as different head and face motion, as given in the LGI city talk session, no algorithm is able to perform reasonable results.

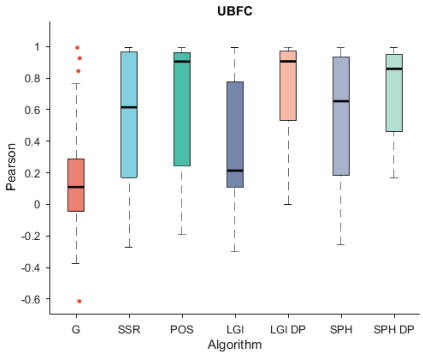
In Figure 5a and 5b the box plot statistics for the UBFC databse are visualized. From Figure 6a and 6b to Figure 9a and 9b the box plot statistics for the different sessions of the LGI database are visualized. In addition to Table 1, the influence of the diffusion process incorporated into the LGI and SPH approach is constituted. Both approaches benefit from its stochastic interpretation as quasi-periodic nature of blood volume changes instead of the classical Fourier based integral approach.

5. Discussion

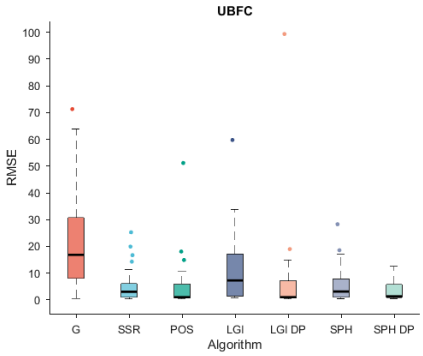
We extended the current knowledge on linear orthogonal operators in the PPGIs feature space. The resulting manifold valued representation is holding implicit properties of invariance with respect to translations of the goup acting onto the set of features. This carries major advantages over the previously prior based assumptions, both POS and LGI, since it comes with a simple change of the topology where the existence of these properties are guaranteed by the fundamental attributes of the space. The computational com-

Database	Green	SSR	POS	LGI	SPH
UBFC	0.16/22.1	0.54/4.95	0.68/4.42	0.75/5.94	0.73/3.21
LGI Resting	0.41/2.61	0.49/1.99	0.41/2.10	0.69/1.41	0.71/1.49
LGI Rotation	0.15/13.2	0.06/10.9	0.12/5.32	0.67/1.92	0.56/2.54
LGI Gym	0.01/33.5	0.03/21.2	0.15/12.2	0.42/2.65	0.26/3.65
LGI Talk	0.15/46.6	0.12/27.8	0.01/37.7	0.51/14.7	0.23/27.8

Table 1: Comparison of heart rate estimation accuracy utilizing different feature operators on diverse face video databases. Each corresponding table entry represents the average Pearson's correlation coefficient together with the average root-mean-square error (RMSE) value in BPM.



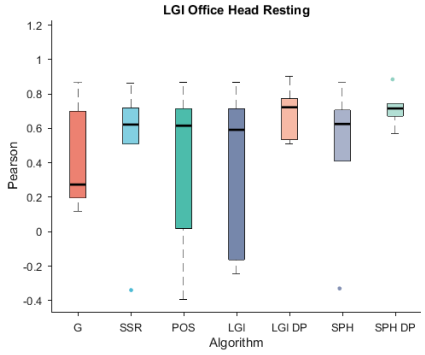
(a) Pearson's correlation coefficients



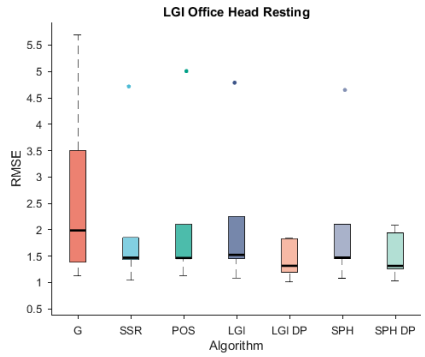
(b) RMSE

Figure 5: UBFC-RPPG: Heart rate estimation accuracy.

plexity of the new feature operator is considerably low and the implementation is fairly easy. The comparison against the most popular representatives of this algorithmic family succeeded with a quite promising strength of prediction accuracy. Since the approach is reflecting a fully closed form solution regarding the original problem statement no nasty tuning of parameters is necessary. It is a parameter free method. The major limitation of the operator in its current form is the restriction of invariance with respect to the group of translations. Since we have presented the proof that the



(a) Pearson's correlation coefficients



(b) RMSE

Figure 6: LGI DB: Heart rate estimation accuracy on Session1: Head Resting.

intensity of pixels does not contribute to the periodic characteristics of skin perfusion, it follows indeed the question concerning the wavelength contribution.

References

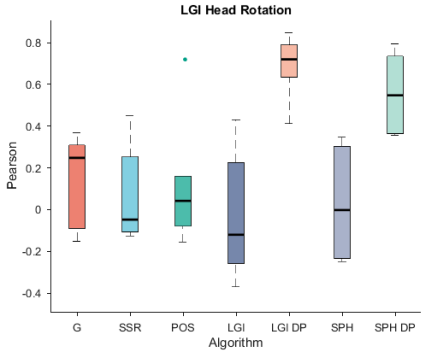
[1] B. Afsari. Riemannian lp center of mass: Existence, uniqueness, and convexity. *Proceedings of the American Mathematical Society.*, 139:655–673, 2011. 5

[2] J. Allen. Photoplethysmography and its application in clinical physiological measurement. *Physiological Measurement*, 28(3):31–39, 2007. 3

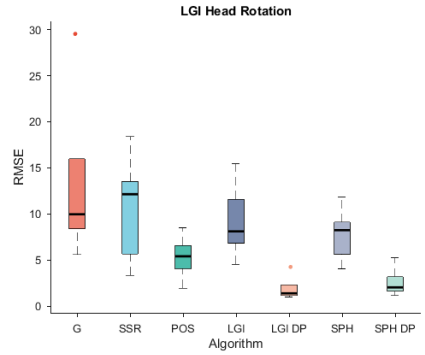
[3] N. Blanik, C. Blazek, C. Pereira, V. Blazek, and S. Leonhardt. Wearable photoplethysmographic sensors: Past and present. *Proc. SPIE 9034, Medical Imaging: Image Processing*, 2014. 3

[4] C. Blazek and M. Hülsbusch. Assessment of allergic skin reactions and their hemodynamical quantification using photoplethysmography imaging. *Computer-aided Noninvasive Vascular Diagnostics. Vol. 3: Proc. of 11th Int. Symposium CNVD*, 3:85–90, 2005. 2

[5] V. Blazek. Optoelektronische Erfassung und rechnerunterstützte Analyse der Mikrozirkulations-Rhythmik. *Biomed. Techn.*, 30(1):121–122, 1985. 2, 3

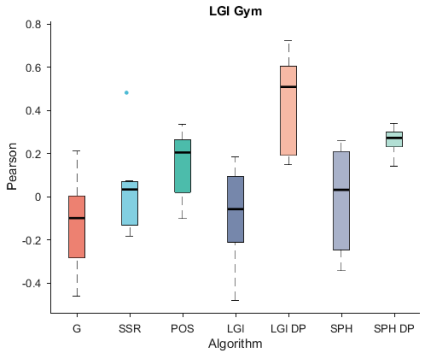


(a) Pearson's correlation coefficients

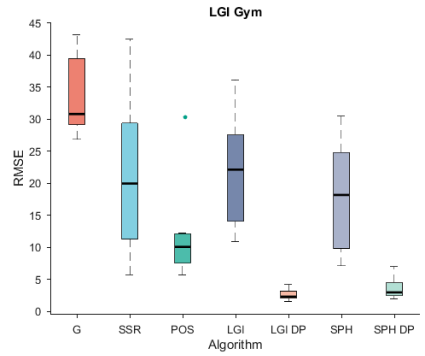


(b) RMSE

Figure 7: LGI DB: Heart rate estimation accuracy on Session2: Head rotation.



(a) Pearson's correlation coefficients



(b) RMSE

Figure 8: LGI DB: Heart rate estimation accuracy on Session3: Bicycle ergometer.

[6] V. Blazek, N. Blanik, C. Blazek, M. Paul, C. Pereira, M. Koeny, B. Venema, and S. Leonhardt. Active and passive optical imaging modality for unobtrusive cardiorespiratory monitoring and facial expressions assessment. *Assessment. Anesth Analg.*, 124:104–119, 2017. 2

[7] S. Bobbia, R. Macwan, Y. Benezeth, A. Mansouri, and J. Dubois. Unsupervised skin tissue segmentation for remote photoplethysmography. *Pattern Recognition Letters*, 2017. 6

[8] J. Cardoso. High-order contrasts for independent component analysis. *Neural Computation*, 11(1):157–192, 1999. 3

[9] W. Chen and D. McDuff. Deepphys: Video-based physiological measurement using convolutional attention networks. *European Conference on Computer Vision (ECCV)*, pages 356–373, 2018. 1, 3, 6

[10] G. de Haan and V. Jeanne. Robust pulse-rate from chrominance-based rppg. *IEEE Transactions on Biomedical Engineering*, 60(10):2878–2886, 2013. 3, 6

[11] G. de Haan and A. van Leest. Improved motion robustness of remote-ppg by using the blood volume pulse signature. *Physiol. Meas*, 3(9):1913–1926, 2014. 3, 6

[12] I. Gelfand. Normierte ringe. *Rech. Math. [Mat. Sbornik] N.S.*, 9:3–24, 1941. 4

[13] A. Henderson, J. Lasselin, M. Lekander, M. Olsson, S. Powis, J. Axelsson, and D. Perrett. Skin colour changes during experimentally-induced sickness. *Brain, Behavior, and Immunity*, 60:312–318, 2017. 2

[14] A. Hertzman. Photoelectric plethysmography of the fingers and toes in man. *Exp. Biol. Med.*, 37(3):529–534, 1937. 2

[15] M. Hülsbusch. A functional imaging technique for optoelectronic assessment of skin perfusion. *PhD thesis, RWTH Aachen University*, 2008. 3, 6

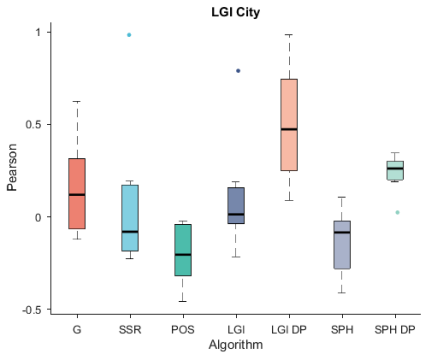
[16] H. Karcher. Riemannian center of mass and mollifier smoothing. *Communications on Pure and Applied Mathematics*, 30:509–541, 1977. 5

[17] X. Li, J. Chen, G. Zhao, and M. Pietikinen. Remote heart rate measurement from face videos under realistic situations. *IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH*, 2014. 1, 3

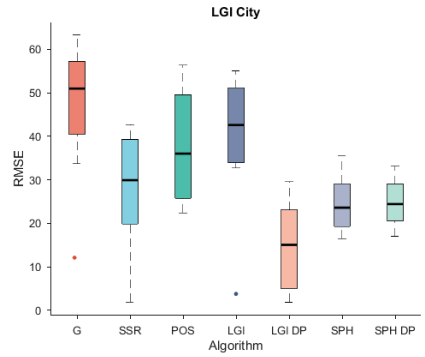
[18] R. Macwan, Y. Benezeth, and A. Mansouri. Remote photoplethysmography with constrained ica using periodicity and chrominance constraints. *BioMedical Engineering OnLine*, 17(1), 2018. 3

[19] K. V. Mardia. Statistics of directional data. *J. Royal Statistical Society. Series B (Methodological)*, 37:197–198, 1975. 5

[20] K. Mathes. Untersuchungen über die Sauerstoffsättigung des menschlichen Arterienblutes. *Arch Exp Path Pharmakol*, (179):698–771, 1935. 2



(a) Pearson's correlation coefficients



(b) RMSE

Figure 9: LGI DB: Heart rate estimation accuracy on Session4: Outdoor city talk.

[21] D. McDuff, E. B. Blackford, and J. R. Estepp. The impact of video compression on remote cardiac pulse measurement using imaging photoplethysmography. *12th IEEE International Conference on Automatic Face and Gesture Recognition*, pages 63–70, 2017. 3

[22] D. McDuff, S. Gontarek, and R. Picard. Remote measurement of cognitive stress via heart rate variability. *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2957–2960, 2014. 2

[23] A. Moço, S. Stuijk, and G. de Haan. Ballistocardiographic artifacts in ppg imaging. *IEEE Transactions on Biomedical Engineering*, 63(9):1804–1811, 2015. 3

[24] A. Moço, S. Stuijk, and G. de Haan. Motion robust ppg-imaging through color channel mapping. *Biomed. Opt. Express*, 7:1737–1754, 2016. 3

[25] H. Molitor and M. Knaizuk. A new bloodless method for continuous recording of peripheral change. *Jour. Phar. Expr. Ther.*, 27:5–16, 1936. 2

[26] A. Osman, J. Turcot, and R. E. Kaliouby. Supervised learning approach to remote heart rate estimation from facial videos. *11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition*, pages 1–6, 2015. 1

[27] C. S. Pilz, J. Krajewski, and V. Blazek. On the diffusion process for heart rate estimation from face videos under realistic conditions. *Pattern Recognition. GCPR 2017. Lecture Notes in Computer Science, vol 10496. Springer*, 10496:361–373, 2017. 1, 3, 4, 6

[28] C. S. Pilz, S. Zaunseder, J. Krajewski, and V. Blazek. Local group invariance for heart rate estimation from face videos in the wild. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 1254–1262, 2018. 1, 3, 6

[29] M. Poh, J. McDuff, and R. Picard. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics Express*, 18(10):1062–10774, 2010. 3

[30] G. Ramirez, O. Fuentes, S.Crites, M. Jimenez, and J. Ordonez. Color analysis of facial skin: Detection of emotional state. *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 474–479, 2014. 2

[31] B. Schölkopf, A. Smola, and V. Vapnik. Prior knowledge in support vector kernels. *Advances in neural information processing systems 10, Cambridge, MA: MIT Press.*, pages 640–646, 1998. 4

[32] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic. A multimodal database for affect recognition and implicit tagging. *IEEE Transactions on Affective Computing*, 3:42–55, 2012. 3

[33] T. Sundelin, M. Lekander, G. Kecklund, E. V. Someren, A. Olsson, and J. Axelsson. Cues of fatigue: Effects of sleep deprivation on facial appearance. *Sleep*, 36(9):1355–1360, 2013. 2

[34] L. Tarassenko, M. Villarroel, A. Guazzi, J. Jorge, D. Clifton, and C. Pugh. Non-contact video-based vital sign monitoring using ambient light and auto-regressive models. *Physiological Measurement*, 35(5):807–831, 2014. 3

[35] S. Tulyakov, X. A. Pineda, E. Ricci, L. Yin, J. Cohn, and N. Sebe. Self-adaptive matrix completion for heart rate estimation from face videos under realistic conditions. *IEEE International Conference on Computer Vision and Pattern Recognition*, 2016. 1, 3

[36] W. Verkruysse, L. Svaasand, and J. Nelson. Remote plethysmographic imaging using ambient light. *Optics Express*, 16(26):21434–21445, 2008. 3, 6

[37] W. Wang, A. den Brinker, S. Stuijk, and G. de Haan. Algorithmic principles of remote ppg. *IEEE Transactions on Biomedical Engineering*, 64(7):1479–1491, 2017. 3, 6

[38] W. Wang, S. Stuijk, and G. de Haan. A novel algorithm for remote photoplethysmography: Spatial subspace rotation. *IEEE Transactions on Biomedical Engineering*, 63(9), 2015. 3, 6

[39] D. Wedekind, A. Trumpp, F. Gaetjen, S. Rasche, K. Matschke, H. Malberg, and S. Zaunseder. Assessment of blind source separation techniques for video-based cardiac pulse extraction. *J. Biomed. Opt.*, 22(3):35002, 2017. 3

[40] N. Wiener. Cybernetics: Or control and communication in the animal and the machine. *MIT Press*, 2(1), 1948. 1