Partial Differential Equations of Bivariate Median Filters

Martin Welk

University for Health Sciences, Medical Informatics and Technology (UMIT),
Eduard-Wallnöfer-Zentrum 1, 6060 Hall/Tyrol, Austria
`martin.welk@umit.at`

Abstract. Multivariate median filters have been proposed as generalisations of the well-established median filter for grey-value images to multi-channel images. As multivariate median, most of the recent approaches use the L^1 median, i.e. the minimiser of an objective function that is the sum of distances to all input points. Many properties of univariate median filters generalise to such a filter. However, the famous result by Guichard and Morel about approximation of the mean curvature motion PDE by median filtering does not have a comparably simple counterpart for L^1 multivariate median filtering. We discuss the affine equivariant Oja median as an alternative to L^1 median filtering. We derive the PDE approximated by Oja median filtering in the bivariate case, and demonstrate its validity by a numerical experiment.

1 Introduction

Median filtering of signals and images goes back to the work of Tukey [17] and has since then been established in image processing as a simple and robust denoising method for grey-value images with favourable structure-preserving properties.

Like other local image filters, the median filter consists of a *selection step* that identifies for each pixel location those pixels which will enter the computation of the filtered value at that location, followed by an *aggregation step* that combines the intensities of these pixels into the filtered value. In the standard setting, the selection step uses a fixed-shape sliding window, which can be called *structuring element* following the naming convention from mathematical morphology. The aggregation step consists in taking the median of the selected intensities. The process can be iterated, giving rise to what is called *iterated median filter.*

The median filter, particularly in its iterated form, has been subject to intensive investigation over the decades. For example, [7] studied so-called *root signals,* non-trivial steady states that occur in the iterated median filter and depend subtly on the choice of the structuring element. Work by Guichard and Morel [9] has identified iterated median filtering as an explicit nonstandard discretisation of (mean) curvature motion [1], thus bridging the discrete filter concept with a partial differential equation (PDE).

Multivariate median filtering. Given the merit of median filtering in processing grey-value images one is interested in stating also a median filter for multi-channel images

such as colour images, flow fields, tensor fields etc. As the switch from single- to multi-channel images does not affect the selection step mentioned above but solely the aggregation, it is clear that what is needed to accomplish this goal is the definition of a multivariate median. A starting point for such a definition is the following characterisation of the univariate median: The median of a finite set of real numbers is the real number that minimises the sum of distances to all numbers of the set. There happens to always exist a number within the given set for which this minimum is attained.

Early attempts to multi-channel median filtering, starting from [2] in 1990, defined therefore a vector-valued median (actually a medoid) that selects *from the set of input points* in $\mathbb{R}^n$ the one that minimises the sum of distances to all other sample points.

More recent approaches, such as [12, 16] for colour images or [22] for symmetric matrices, rely on the same minimisation but without the restriction to the given data points. The underlying multivariate median concept is known in the statistics literature as *spatial median* or L^1 *median*. It can be traced back to work by Hayford from 1902 [10] and Weber from 1909 [18], followed by [3, 8, 19] and many others.

However, this is not the only multivariate median concept in literature. For example, the *simplex median* established by Oja in 1983 [13] generalises the distances between points on the real line from the univariate median definition not into distances but into simplex volumes in higher dimensions. Thus, the simplex median of a finite set of points in $\mathbb{R}^n$ is the point $\boldsymbol{p} \in \mathbb{R}^n$ that minimises the sum of simplex volumes $|[\boldsymbol{p}, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n]|$ where $\boldsymbol{a}_i$ are distinct points of the input data set. An advantage of this concept that is relevant for many statistics applications is its affine equivariance, i.e. that it commutes with affine transformations of the data space. In contrast, the L^1 median only affords Euclidean equivariance. The Oja simplex median is not the only affine equivariant median concept; other concepts have been developed by modifying the L^1 median e.g. in [4, 11, 14], see also the survey in [5]. For further multivariate concepts see the review [15].

Multivariate median filters and PDE. While the above-mentioned relationship between univariate median filtering and the mean curvature motion PDE could be extended to relate also adaptive median filtering procedures [21] and further discrete filters [20] to well-understood PDEs of image processing, the picture changes when turning to multivariate median filtering. As demonstrated in [20], it is possible to derive some PDE for median filtering based on the spatial median as in [16]. However, this PDE involves complicated coefficient functions coming from elliptic integrals most of which cannot even be stated in closed form, see Section 3.1 of this paper.

Therefore the question arises whether other multivariate median concepts could be advantageous in multi-channel image processing. The present paper is intended as a first step in this direction.

Our contribution. In this paper, we focus strictly on the bivariate case (e.g. two-channel images or 2D flow fields). We juxtapose the L^1 median and Oja median in the context of image filtering as well as in terms of basic geometric properties, and present PDEs approximated by both kinds of median filters. The novel PDE for Oja median filtering is validated by a numerical experiment.
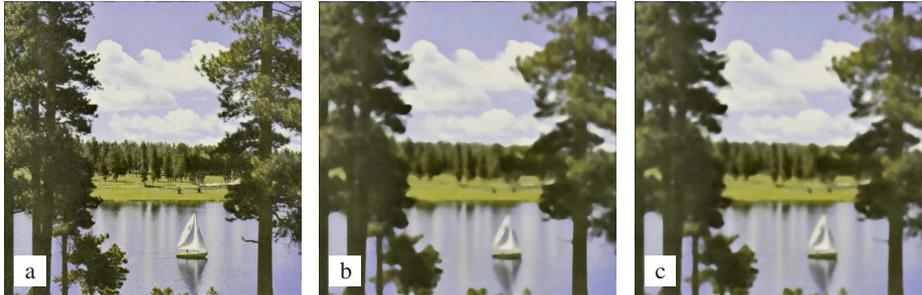
Fig. 1. Median filtering of an image with two colour channels. **(a)** Test image, 512×512 pixels, reduced to a yellow–blue colour space. – **(b)** Filtered by one L^1 median filtering step with a disc-shaped structuring element of radius 5. – **(c)** Filtered by one Oja median filtering step with the same structuring element as in (b).

Structure of the paper. In Section 2, we demonstrate bivariate median filtering by L^1 and Oja median on a simple two-channel colour image, and discuss basic geometric properties of both median concepts. Section 3 is dedicated to PDE approximation results that generalise Guichard and Morel's [9] result for univariate median filtering. For L^1 median filtering, the known result from [20] is rephrased more explicitly for the bivariate case. For Oja median filtering, a PDE is derived in Section 3.2 for the first time, which is the main result of this paper. This PDE is afterwards validated by a numerical experiment in Section 4. A summary and outlook in Section 5 conclude the paper.

2   Comparison of L^1 and Oja Median

Median filter demonstration. We start by demonstrating L^1 and Oja medians in the role of image filters. Since this paper focusses on the bivariate case, we can think e.g. of two-channel colour images or 2D flow fields as examples. While the latter are practically more relevant than the earlier, we prefer two-channel images here because the main focus of the present paper is theoretical, and two-channel images are visually easy to understand. To this end, a RGB colour image has been reduced to a yellow–blue colour space by averaging the red and green channels, see Figure 1(a). This image has been filtered by L^1 and Oja median filtering with identical parameters in Figure 1(b) and (c), respectively. The results of both filters look fairly similar. They display the same kind of structure simplification and smoothing contours as known from median filters.

Geometric facts about L^1 and Oja median. To add some geometric intuition about the bivariate median filters under investigation, we consider small point sets in the plane and their medians. The following statements can easily be inferred from standard elementary geometry arguments such as the triangle inequality (for the L^1 median) and multiplicities of covering of the convex hull of input points by the triangles with input and median points as corners (for the 2D Oja median).

For two points, the L^1 median criterion is fulfilled equally for all points of their connecting line segment. The Oja median criterion is even fulfilled by all points of
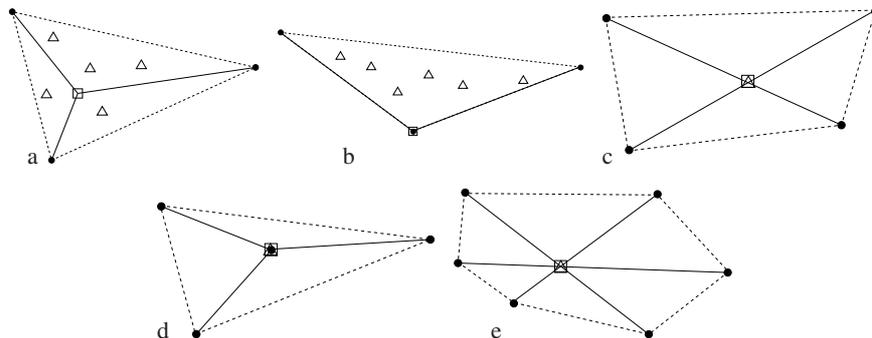
Fig. 2. Simple configurations of input data points (solid points) with their L^1 medians (squares) and Oja medians (triangles). **(a)** Three points forming a triangle with all interior angles less than 120 degrees: The L^1 median is the *Steiner point*; any point within the triangle is an Oja median. – **(b)** Three points forming a triangle with an obtuse angle of 120 degrees or more: The obtuse corner is the L^1 median; still, all points within the triangle are Oja medians. – **(c)** Four points forming a convex quadrangle: L^1 and Oja median coincide at the intersection of the diagonals. – **(d)** Four points whose convex hull is a triangle: L^1 and Oja median coincide at the data point that is not a corner of the convex hull. – **(e)** $2n$ points that form a convex $2n$-gon (hexagon shown as example) in which all diagonals between opposing points have a common intersection point: L^1 and Oja median coincide at this intersection point.

the straight line through these points since the Oja median definition degenerates for collinear sets of points. While this degeneracy can be mitigated by adding a continuity criterion, we do not further treat degenerate cases within the present paper.

For three points, the L^1 median depends on the sort of triangle they span. If all of its interior angles are smaller than 120 degrees, see Figure 2(a), the sum of distances to the corners is minimised by a unique point known as *Steiner point* or *Fermat-Torricelli point,* from which all sides of the triangle are seen under 120 degree angles. For a triangle with an obtuse corner of at least 120 degrees, this corner is the L^1 median, see Figure 2(b). The Oja median criterion is met in both cases by all points of the triangle. This is consistent with the affine equivariance of the Oja median that does not discriminate triangles by shape. It also shows how indeed simplices take the role of line segments from the univariate median definition – the three-point case of the bivariate Oja median is the analogue of the two-point case of the univariate median.

For four points, L^1 and Oja median always coincide: If the convex hull of the data points is a triangle, then the data point that is not a corner of the convex hull is the median, see Figure 2(d); if it is a convex quadrangle, then the intersection point of its diagonals is the median, see Figure 2(c).

The coincidence between L^1 and Oja median continues also in some configurations of more data points. A (non-generic) example is shown in Figure 2(e): A convex $2n$-gon in which all the diagonals that bisect the point set (i.e. those that span n sides) have a common intersection point, features this point as L^1 and Oja median.

We point out two observations that can be made from these simple configurations. Firstly, bivariate medians, unlike their univariate counterpart, cannot always be chosen

from the input data set, but they happen to be input data points in some generic configurations. Only in cases when none of the input points lies sufficiently "in the middle" of the data, a new point is created. Secondly, despite their different definitions, also L^1 and Oja median coincide in some generic situations, or are not far apart from each other. This adds plausibility to why the image filtering results in Figure 1 are that similar.

3 PDE Limit Analysis

In this section, we study median filters in a space-continuous setting. As proven in [9], a univariate median filtering step of an image with disc-shaped structuring element of radius ϱ approximates for $\varrho \to 0$ a time step of size $\tau = \varrho^2/6$ of an explicit scheme for the mean curvature motion PDE. We will present PDEs that are approximated in the same sense by L^1 and Oja median filtering of bivariate images.

The reformulation of a local image filter to a space-continuous setting is straightforward. The main modification is that the set of values that is returned by the selection step and is processed further in the aggregation step is now infinite and equipped with a density. This density is induced from the uniform distribution of function arguments in the structuring element in the image domain via the Jacobian of the image function.

3.1 L^1 Median

An analysis of the L^1 multivariate median filter for images $\boldsymbol{u} : \mathbb{R}^2 \supset \Omega \to \mathbb{R}^n$ has been given in [20]. Here, we break down the essential result to the bivariate case $n = 2$.

Proposition 1. *One step of L^1 median filtering of a bivariate image $\boldsymbol{u} : \mathbb{R}^2 \supset \Omega \to \mathbb{R}^2$, $(x, y) \mapsto (u, v)$, with a disc-shaped structuring element D_ϱ of radius ϱ approximates for $\varrho \to 0$ an explicit time step of size $\tau = \varrho^2/6$ of the PDE system*

$$\begin{pmatrix} u_t \\ v_t \end{pmatrix} = \boldsymbol{S}(\boldsymbol{\nabla} u, \boldsymbol{\nabla} v) \begin{pmatrix} u_{\boldsymbol{\eta\eta}} \\ v_{\boldsymbol{\eta\eta}} \end{pmatrix} + \boldsymbol{T}(\boldsymbol{\nabla} u, \boldsymbol{\nabla} v) \begin{pmatrix} u_{\boldsymbol{\xi\xi}} \\ v_{\boldsymbol{\xi\xi}} \end{pmatrix} \tag{1}$$

where $\boldsymbol{\eta}$ is the major, and $\boldsymbol{\xi}$ the minor eigenvector of the structure tensor $\boldsymbol{J} := \boldsymbol{J}(\boldsymbol{\nabla} u, \boldsymbol{\nabla} v) := \boldsymbol{\nabla} u \boldsymbol{\nabla} u^{\mathrm{T}} + \boldsymbol{\nabla} v \boldsymbol{\nabla} v^{\mathrm{T}}$. The coefficient matrices $\boldsymbol{S}(\boldsymbol{\nabla} u, \boldsymbol{\nabla} v)$, $\boldsymbol{T}(\boldsymbol{\nabla} u, \boldsymbol{\nabla} v)$ are given by

$$\boldsymbol{S}(\boldsymbol{\nabla} u, \boldsymbol{\nabla} v) := \boldsymbol{R} \operatorname{diag}\!\Big(Q_1\big(|\partial_{\boldsymbol{\eta}} \boldsymbol{u}|/|\partial_{\boldsymbol{\xi}} \boldsymbol{u}|\big), Q_2\big(|\partial_{\boldsymbol{\eta}} \boldsymbol{u}|/|\partial_{\boldsymbol{\xi}} \boldsymbol{u}|\big)\Big) \boldsymbol{R}^{\mathrm{T}} , \tag{2}$$

$$\boldsymbol{T}(\boldsymbol{\nabla} u, \boldsymbol{\nabla} v) := \boldsymbol{R} \operatorname{diag}\!\Big(Q_2\big(|\partial_{\boldsymbol{\xi}} \boldsymbol{u}|/|\partial_{\boldsymbol{\eta}} \boldsymbol{u}|\big), Q_1\big(|\partial_{\boldsymbol{\xi}} \boldsymbol{u}|/|\partial_{\boldsymbol{\eta}} \boldsymbol{u}|\big)\Big) \boldsymbol{R}^{\mathrm{T}} , \tag{3}$$

where $\boldsymbol{R} = (\mathrm{D}\boldsymbol{u}^{-1})^{\mathrm{T}} \boldsymbol{P} \operatorname{diag}(|\partial_{\boldsymbol{\eta}} \boldsymbol{u}|, |\partial_{\boldsymbol{\xi}} \boldsymbol{u}|)$ is a rotation matrix that depends on the Jacobian $\mathrm{D}\boldsymbol{u}$ of $\boldsymbol{u}$ and the eigenvector matrix $\boldsymbol{P} = (\boldsymbol{\eta} \,|\, \boldsymbol{\xi})$ of $\boldsymbol{J}$. The functions $Q_1, Q_2 : [0, \infty] \to \mathbb{R}$ are given by the quotients of elliptic integrals

$$Q_1(\lambda) = 3 \iint\limits_{D_1(\boldsymbol{0})} \frac{s^2 t^2}{(s^2 + \lambda^2 t^2)^{3/2}} \, \mathrm{d}s \, \mathrm{d}t \Big/ \iint\limits_{D_1(\boldsymbol{0})} \frac{s^2}{(s^2 + \lambda^2 t^2)^{3/2}} \, \mathrm{d}s \, \mathrm{d}t , \tag{4}$$

$$Q_2(\lambda) = 3 \iint\limits_{D_1(\boldsymbol{0})} \frac{t^4}{(s^2 + \lambda^2 t^2)^{3/2}} \, \mathrm{d}s \, \mathrm{d}t \Big/ \iint\limits_{D_1(\boldsymbol{0})} \frac{t^2}{(s^2 + \lambda^2 t^2)^{3/2}} \, \mathrm{d}s \, \mathrm{d}t \tag{5}$$

for $\lambda \in (0, \infty)$, together with the limits $Q_1(0) = Q_2(0) = 1$, $Q_1(\infty) = Q_2(\infty) = 0$.

The proof relies on the following statement which is proven in [20].

Lemma 1 (from [20]). *Let $\boldsymbol{u}$ be given as in Proposition 1. Assume that the Jacobian $\mathrm{D}\boldsymbol{u}$ at some location (x, y) is diagonal, i.e. $u_y = v_x = 0$, and $u_x \geq v_y \geq 0$. Then one step of L^1 median filtering with structuring element D_ϱ approximates for $\varrho \to 0$ at (x, y) an explicit time step of size $\tau = \varrho^2/6$ of the PDE system*

$$
\begin{aligned}
u_t &= Q_1(u_x/v_y)u_{xx} + Q_2(v_y/u_x)u_{yy} \; , \\
v_t &= Q_2(u_x/v_y)v_{xx} + Q_1(v_y/u_x)v_{yy} \; ,
\end{aligned}
\tag{6}
$$

with the coefficient functions Q_1, Q_2 as stated in Proposition 1.

Proof (of Proposition 1). Consider an arbitrary fixed location (x^*, y^*). By applying rotations with $\boldsymbol{P}$ in the x-y plane and with $\boldsymbol{R}$ in the u-v plane, x, y can be aligned with the eigenvector directions $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$, and u, v with the corresponding derivatives $\partial_{\boldsymbol{\eta}}\boldsymbol{u}$, $\partial_{\boldsymbol{\xi}}\boldsymbol{u}$. Then Lemma 1 can be applied. Reverting the rotations in the x-y and u-v planes, the PDE system (6) turns into the system (1)–(3) of the proposition. $\square$

Remark 1. The derivation of the PDE (1) from a special case by Euclidean transform immediately implies its Euclidean equivariance.

Remark 2. The vectors $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ used in (1)–(3) are the directions of greatest and least change of the bivariate function $\boldsymbol{u}$, thus the closest analoga to gradient and level line directions of univariate images, see [6]. The use of these image-adaptive local coordinates characterises (1) as a curvature-based PDE remotely similar to the (mean) curvature motion PDE approximated by univariate median filtering.

3.2   Oja's Simplex Median

Theorem 1. *Let a bivariate image $\boldsymbol{u} : \mathbb{R}^2 \supset \Omega \to \mathbb{R}^2$, $(x, y) \mapsto (u, v)$, be given. At any location where $\det \mathrm{D}\boldsymbol{u} \neq 0$, one step of Oja median filtering of $\boldsymbol{u}$ with the structuring element D_ϱ approximates for $\varrho \to 0$ an explicit time step of size $\tau = \varrho^2/24$ of the PDE system*

$$
\begin{pmatrix} u_t \\ v_t \end{pmatrix} = 2 \begin{pmatrix} u_{xx} + u_{yy} \\ v_{xx} + v_{yy} \end{pmatrix} - \boldsymbol{A}(\boldsymbol{\nabla}u, \boldsymbol{\nabla}v) \begin{pmatrix} u_{xx} - u_{yy} \\ v_{yy} - v_{xx} \end{pmatrix} + \boldsymbol{B}(\boldsymbol{\nabla}u, \boldsymbol{\nabla}v) \begin{pmatrix} u_{xy} \\ -v_{xy} \end{pmatrix} \tag{7}
$$

with the coefficient matrices

$$
\boldsymbol{A}(\boldsymbol{\nabla}u, \boldsymbol{\nabla}v) := \frac{1}{u_x v_y - u_y v_x} \begin{pmatrix} u_x v_y + u_y v_x & 2u_x u_y \\ 2v_x v_y & u_x v_y + u_y v_x \end{pmatrix} \; , \tag{8}
$$

$$
\boldsymbol{B}(\boldsymbol{\nabla}u, \boldsymbol{\nabla}v) := \frac{2}{u_x v_y - u_y v_x} \begin{pmatrix} u_x v_x - u_y v_y & u_x^2 - u_y^2 \\ v_x^2 - v_y^2 & u_x v_x - u_y v_y \end{pmatrix} \; . \tag{9}
$$
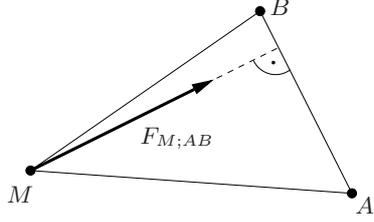
Fig. 3. Anti-gradient vector $F_{M;AB}$ for the area of a triangle MAB with variable point M.

Proof. We consider the median of the values $\boldsymbol{u}(x,y)$ within the Euclidean ϱ-neighbourhood of $(0,0)$, and assume $\det \mathrm{D}\boldsymbol{u}(\boldsymbol{0}) \neq 0$. By the affine equivariance of Oja's simplex median, the u-v plane can be transformed in such a way that the Jacobian $\mathrm{D}\boldsymbol{u}$ at $(x,y) = (0,0)$ becomes equal to the unit matrix, i.e. $u_x = v_y = 1$, $u_y = v_x = 0$, and $u(0,0) = v(0,0) = 0$. Then the Taylor expansion of (u,v) up to second order around $(0,0)$ reads as

$$\begin{pmatrix} u(x,y) \\ v(x,y) \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \alpha x^2 + \beta xy + \gamma y^2 \\ \delta x^2 + \varepsilon xy + \zeta y^2 \end{pmatrix} , \tag{10}$$

where the coefficients are given by derivatives of u, v at $(x,y) = (0,0)$ as

$$\alpha = \tfrac{1}{2}u_{xx}(0,0) , \qquad \beta = u_{xy}(0,0) , \qquad \gamma = \tfrac{1}{2}u_{yy}(0,0) , \tag{11}$$
$$\delta = \tfrac{1}{2}v_{xx}(0,0) , \qquad \varepsilon = v_{xy}(0,0) , \qquad \zeta = \tfrac{1}{2}v_{yy}(0,0) . \tag{12}$$

Restating the definition of Oja's simplex median for continuous data sets with density function $f(u,v)$, we seek the point $M := (u^*, v^*)$ which minimises the integral over all areas of triangles MAB with $A = (u_1, v_1)$ and $B = (u_2, v_2)$ with $(u_1, v_1) = (u(x_1, y_1), v(x_1, y_1))$, $(u_2, v_2) = (u(x_2, y_2), v(x_2, y_2))$, $(x_1, y_1), (x_2, y_2) \in D_\varrho(0,0)$, weighted with the density $f(u_1, v_1)f(u_2, v_2)$.

For each triangle MAB, the negative gradient of its area as function of M is a force vector $F_{M;AB}$ perpendicular to AB with a length proportional to the length $|AB|$, see Figure 3. Assuming that MAB is positively oriented, this vector equals $\tfrac{1}{2}(y_2 - y_1, -x_2 + x_1)$.

Sorting the pairs (A, B) by the orientation angles φ of the lines AB, we see that the minimisation condition for the Oja median can be expressed as

$$\boldsymbol{\Phi}(u^*, v^*) = \frac{1}{2} \int_0^{2\pi} \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix} F(u^*, v^*, \varphi) \, \mathrm{d}\varphi = 0 . \tag{13}$$

Here, $F(\varphi)$ is essentially the resultant of all forces $F_{M;AB}$ for which the line AB intersects the ray from M in direction $(\cos\varphi, \sin\varphi)$ perpendicularly. Each force $F_{M;AB}$ is weighted with the combined density $f(A)f(B) = f(u_1, v_1)f(u_2, v_2)$.

Moreover, u^*, v^* will be of order $\mathcal{O}(\varrho)$ (in fact, even $\mathcal{O}(\varrho^2)$). Thus, (u^*, v^*) can be expressed up to higher order terms via linearisation as

$$\begin{pmatrix} u^* \\ v^* \end{pmatrix} = -\big(\mathrm{D}\boldsymbol{\Phi}(0,0)\big)^{-1}\boldsymbol{\Phi}(0,0) . \tag{14}$$

We therefore turn now to derive an expression for $F(0, 0, \varphi)$. Considering first $\varphi = 0$, this means that all point pairs (A, B) in the u-v right half-plane with $u_1 = u_2$ contribute to $F(0, 0, 0)$, yielding

$$F(0, 0, 0) = \int_0^{+\infty} \int_{-\infty}^{+\infty} \int_v^{+\infty} f(u, v)f(u, w)(w - v)^2 \, \mathrm{d}w \, \mathrm{d}v \, \mathrm{d}u \ . \qquad (15)$$

Note that the factor $(w - v)$ occurs squared in the integrand. One factor $(w - v)$ originates from the length of the triangle baseline AB. The second factor $(w - v)$ results from the fact that we have organised in (13), (15) an integration over point pairs (A, B) in the plane using a polar coordinate system similar to a Radon transform; $w - v$ arises as the Jacobian of the corresponding coordinate transform from Cartesian to Radon coordinates. The derivatives of $F(u^*, v^*, 0)$ with regard to the coordinates of M are

$$F_{u^*}(0, 0, 0) = -\int_{-\infty}^{+\infty} \int_v^{+\infty} f(0, v)f(0, w)(w - v)^2 \, \mathrm{d}w \, \mathrm{d}v \ , \qquad (16)$$

$$F_{v^*}(0, 0, 0) = 0 \ . \qquad (17)$$

Forces $F(0, 0, \varphi)$ and their derivatives for arbitrary angles φ can be obtained from (15), (16), (17) by rotating the u, v coordinates accordingly.

When considering a ϱ-neighbourhood of $(x, y) = (0, 0)$, the density $f(u, v)$ is zero outside of an $\mathcal{O}(\varrho)$-neighbourhood of $(0, 0)$, allowing to limit the indefinite integrals from (15) to the intervals $u \in [u^*, \bar{u}]$, $v \in [\underline{v}(u), \bar{v}(u)]$ and $w \in [v, \bar{v}(u)]$ such that

$$F(0, 0, 0) = \int_0^{\bar{u}} \int_{\underline{v}(u)}^{\bar{v}(u)} \int_v^{\bar{v}(u)} f(u, v)f(u, w)(w - v)^2 \, \mathrm{d}w \, \mathrm{d}v \, \mathrm{d}u \qquad (18)$$

and similarly for (16).

To compute $F(0, 0, 0)$ and $F_{u^*}(0, 0, 0)$, we write them as functions of the coefficients of (10), i.e. $F(0, 0, 0) =: G(\alpha, \beta, \gamma, \delta, \varepsilon, \zeta)$ and $F_{u^*}(0, 0, 0) =: H(\alpha, \beta, \gamma, \delta, \varepsilon, \zeta)$.

We will linearise G and H around the point $(\alpha, \beta, \gamma, \delta, \varepsilon, \zeta) = \mathbf{0}$ that represents the linear function $(u(x, y), v(x, y)) = (x, y)$. To justify this linearisation, remember that we are interested in the limit $\varrho \to 0$, such that only the terms of lowest order in ϱ matter. Cross-effects between the different coefficients occur only in higher order terms. Denoting from now on by $\doteq$ equality up to higher order terms, we have therefore

$$G \doteq G^0 + G_\alpha^0 \alpha + G_\beta^0 \beta + G_\gamma^0 \gamma + G_\delta^0 \delta + G_\varepsilon^0 \varepsilon + G_\zeta^0 \zeta \ , \qquad (19)$$

$$H \doteq H^0 + H_\alpha^0 \alpha + H_\beta^0 \beta + H_\gamma^0 \gamma + H_\delta^0 \delta + H_\varepsilon^0 \varepsilon + H_\zeta^0 \zeta \qquad (20)$$

where G^0, G_α^0 etc. are short for $G(\mathbf{0})$, $G_\alpha(\mathbf{0})$ etc.

To compute G^0 and H^0, we insert into (15) the bounds $\bar{u} = \varrho$, $\bar{v}(u) = \sqrt{\varrho^2 - u^2}$, $\underline{v}(u) = -\bar{v}(u)$. The density becomes constant within the region defined by $\bar{u}$, $\underline{v}(u)$ and $\bar{v}(u)$, with $f(u, v) = 1$. Thus (18) and (16) yield

$$G^0 = \tfrac{32}{45} \varrho^5 \ , \qquad\qquad H^0 = -\tfrac{4}{3} \varrho^4 \ . \qquad (21)$$

For G_α^0 and H_α^0, one has to vary α to obtain the bounds $\bar{u} = \varrho + \alpha\varrho^2$, $\bar{v}(u) = \sqrt{\varrho^2 - u^2 - 2\alpha u^3}$, $\underline{v}(u) \doteq -\bar{v}(u)$. The density $f(u, v)$ within the so-given bounds

is $1/\det(\mathrm{D}\boldsymbol{u})$ at the location $(x(u,v), y(u,v))$ with $x = u - \alpha u^2 + \mathcal{O}(\varrho^3)$, $y = v$, i.e. $f(u,v) = 1 - 2\alpha u + \mathcal{O}(\varrho^2)$. Thus we have

$$G_\alpha^0 \doteq \frac{\mathrm{d}}{\mathrm{d}\alpha} \int_0^{\bar{u}} \int_{\underline{v}(u)}^{\bar{v}(u)} \int_v^{\bar{v}(u)} (1 - 2\alpha u)^2 (w - v)^2 \,\mathrm{d}w \,\mathrm{d}v \,\mathrm{d}u \bigg|_{\alpha=0} = -\tfrac{4}{9}\varrho^6 \,, \quad (22)$$

$$H_\alpha^0 \doteq -\frac{\mathrm{d}}{\mathrm{d}\alpha} \int_{\underline{v}(0)}^{\bar{v}(0)} \int_v^{\bar{v}(0)} (w - v)^2 \,\mathrm{d}w \,\mathrm{d}v \bigg|_{\alpha=0} = 0 \,. \tag{23}$$

Proceeding similarly for the other coefficients, we find

- for G_β^0, H_β^0: $\bar{u} = \varrho$, $\bar{v} \doteq \sqrt{\varrho^2 - u^2 + \beta^2 u^4} + \beta u^2$, $\underline{v} \doteq -\sqrt{\varrho^2 - u^2 + \beta^2 u^4} + \beta u^2$, $f(u,v) \doteq 1 - \beta v$;
- for G_γ^0, H_γ^0: $\bar{u} = \varrho$, $\bar{v} \doteq \sqrt{(\varrho^2 - u^2)(1 + 2\gamma u)}$, $\underline{v} \doteq -\bar{v}$, $f(u,v) = 1$;
- for G_δ^0, H_δ^0: $\bar{u} = \varrho$, $\bar{v} \doteq \sqrt{\varrho^2 - u^2 + \delta^2 u^4} + \delta u^2$, $\underline{v} \doteq -\sqrt{\varrho^2 - u^2 + \delta^2 u^4} + \delta u^2$, $f(u,v) = 1$;
- for $G_\varepsilon^0, H_\varepsilon^0$: $\bar{u} = \varrho$, $\bar{v} \doteq \sqrt{(\varrho^2 - u^2)(1 + 2\varepsilon u)}$, $\underline{v} \doteq -\bar{v}$, $f(u,v) \doteq 1 - \varepsilon u$;
- for G_ζ^0, H_ζ^0: $\bar{u} = \varrho$, $\bar{v} \doteq \sqrt{\varrho^2 - u^2 + \zeta(\varrho^2 - u^2)}$, $\underline{v} \doteq -\sqrt{\varrho^2 - u^2 + \zeta(\varrho^2 - u^2)}$, $f(u,v) \doteq 1 - 2\zeta v$.

From these it follows that

$$G_\gamma^0 = \tfrac{8}{9}\varrho^6 \,, \quad G_\varepsilon^0 = \tfrac{4}{9}\varrho^6 \,, \quad G_\beta^0 = G_\delta^0 = G_\zeta^0 = H_\beta^0 = H_\gamma^0 = H_\delta^0 = H_\varepsilon^0 = H_\zeta^0 = 0 \,. \tag{24}$$

Inserting (21), (22), (23), (24) into (19) and (20), we have

$$F(0,0,0) = \tfrac{32}{45}\varrho^5 + \tfrac{4}{9}\varrho^6(-\alpha + 2\gamma + \varepsilon) \,, \qquad F_{u^*}(0,0,0) = \tfrac{4}{3}\varrho^4 \,, \tag{25}$$

and by orthogonal transform in the u-v plane

$$
\begin{aligned}
F(0,0,\varphi) = \tfrac{32}{45}\varrho^5 + \tfrac{4}{9}\varrho^6 \Big(&-(\alpha\cos\varphi + \delta\sin\varphi)\cos^2\varphi \\
&- 2(\beta\cos\varphi + \varepsilon\sin\varphi)\cos\varphi\sin\varphi - (\gamma\cos\varphi + \zeta\sin\varphi)\sin^2\varphi \\
&+ 2(\alpha\cos\varphi + \delta\sin\varphi)\sin^2\varphi - 4(\beta\cos\varphi + \varepsilon\sin\varphi)\cos\varphi\sin\varphi \\
&+ 2(\gamma\cos\varphi + \zeta\sin\varphi)\cos^2\varphi - 2(-\alpha\sin\varphi + \delta\cos\varphi)\cos\varphi\sin\varphi \\
&+ (-\beta\sin\varphi + \varepsilon\cos\varphi)(\cos^2\varphi - \sin^2\varphi) \\
&+ 2(-\gamma\sin\varphi + \zeta\cos\varphi)\cos\varphi\sin\varphi \Big) \,,
\end{aligned}
\tag{26}
$$

$$F_{u^*}(0,0,\varphi) = \tfrac{4}{3}\varrho^4 \cos\varphi \,, \quad F_{v^*}(0,0,\varphi) = \tfrac{4}{3}\varrho^4 \sin\varphi \,. \tag{27}$$

Integration (13) then yields

$$\boldsymbol{\Phi}(0,0) = \frac{\pi}{18}\varrho^6 \begin{pmatrix} \alpha + 3\gamma - \varepsilon \\ -\beta + 3\delta - \zeta \end{pmatrix} \,, \qquad \mathrm{D}\boldsymbol{\Phi}(0,0) = -\frac{2}{3}\pi\varrho^4 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{28}$$

and via (14) eventually

$$\begin{pmatrix} u^* \\ v^* \end{pmatrix} = \frac{\varrho^2}{12} \begin{pmatrix} \alpha + 3\gamma - \varepsilon \\ -\beta + 3\delta + \zeta \end{pmatrix} \,. \tag{29}$$

Inserting (11), (12) into (29), we see that for $\mathrm{D}\boldsymbol{u} = \mathrm{diag}(1,1)$ the Oja median filtering step approximates an explicit time step of size $\tau = \varrho^2/24$ of the PDE system

$$u_t = u_{xx} + 3u_{yy} - 2v_{xy} , \qquad\qquad v_t = 3v_{xx} + v_{yy} - 2u_{xy} . \qquad (30)$$

Transfer to the general case with arbitrary $\mathrm{D}\boldsymbol{u}$ is accomplished by an affine transformation in the u-v coordinates. This yields the PDE system from equations (7)–(9) of the theorem. $\qquad\square$

Remark 3. The derivation of the PDE of Theorem 1 by affine transformation immediately implies its affine equivariance.

Remark 4. The equations (7)–(9) are degenerate at locations where $\det \mathrm{D}\boldsymbol{u} = 0$. This corresponds to the degeneracy of the Oja median itself for collinear input data. Future work will be concerned with this non-generic case.

Remark 5. The eigenvector directions $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ of the structure tensor do not appear in a natural way in the presentation of (7). This is plausible because these eigenvectors are strongly related with a Euclidean geometry concept of the u-v plane, and are thereby inappropriate for an affine equivariant process like Oja median filtering.

4   Experimental Demonstration of Oja Median Filtering

To demonstrate the validity of the PDE approximation result stated in Theorem 1, we consider a simple bivariate example function given by

$$u(x,y) = x^2 , \qquad\qquad v(x,y) = \sqrt{x^2 + y^2} . \qquad (31)$$

Level sets of u and v for this function in the range $[0,1] \times [0,1]$ are depicted in Figure 4. In this figure, also seven test locations a–g are depicted for which Table 1 contains analytically computed time steps $(\tau u_1, \tau v_t)$ of the PDE (7) and results of numerical approximations of the Oja median on a discrete grid within a structuring element of radius $\varrho = 0.1$. For the Oja median computation, we first normalise the input data set by a principal axis transform to zero mean and isotropic unit standard deviation, and apply then a gradient descent method with adaptive step size control, after which the normalising transform is reverted.

The results for locations b, c, d and g in Table 1 confirm that in the generic setting where $\det \mathrm{D}\boldsymbol{u}$ is sufficiently far away from zero, the PDE time step and the median update $(u^* - u, v^* - v)$ computed on the discrete grid match each other well. The observed relative errors in the range of $\leq 3\%$ (based on the Euclidean norm of $(\tau u_t, \tau v_t)$) are reasonable given the structuring element radius $\varrho = 0.1$ and the grid resolution.

Larger discrepancies are observed for locations a, e, and f which are closer to the coordinate axes. Note that on the x axis, $\mathrm{D}\boldsymbol{u}$ becomes singular due to coinciding gradient directions for u and v, while on the y axis it does so due to the vanishing of $\boldsymbol{\nabla} u$. These discrepancies indicate numerical problems of the median computation rather than inaccuracy of the PDE, pointing to the need for algorithmic improvements in the Oja median computation.
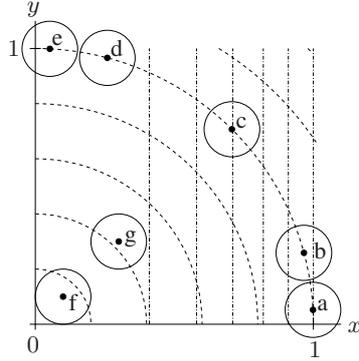
Fig. 4. Example function $(u, v) = (x^2, \sqrt{x^2 + y^2})$ used to demonstrate PDE approximation of Oja median filtering. Dot-dashed lines are level lines $u = \mathrm{const}$, dashed lines are level lines $v = \mathrm{const}$. Points a–g are the sample locations for which numerical results are given in Table 1, surrounded by their structuring elements as solid circles.

5   Summary and Outlook

In this paper, we have analysed the Oja median filter in the bivariate case and shown that it asymptotically approximates for vanishing structuring element size a second-order PDE which is more favourable than that approximated by the more popular L^1 median filter. This gives a strong motivation to deeper investigate the applicability of Oja median filtering in multi-channel image processing.

The proof of Theorem 1 will be presented in more detail in forthcoming work, where also the degenerate case $\det \mathrm{D}\boldsymbol{u}$ will be given broader attention. It will be of high interest to extend our results from the bivariate to the general multivariate case. Moreover, analysis of further multivariate median filters proposed in the statistical literature, as mentioned in the introduction, is another goal of ongoing research.

Finally, a straightforward computation of the Oja median filter has higher complexity than that of the L^1 median filter (due to the point pairs to be iterated over instead of points). Further development of efficient algorithms will therefore be important in order to establish Oja median filtering as a practical image filter.

Table 1. Comparison of analytically computed time steps $(\tau u_t, \tau v_t)$ of the PDE (7)–(9) with numerical computation of the Oja median (u^*, v^*) for the function $(u, v) = (x^2, \sqrt{x^2 + y^2})$. To compute (u^*, v^*), the structuring element of radius $\varrho = 0.1$ was sampled using a grid with spatial mesh size $h = 0.001$, generating about 31,000 data points. The time step size for (7) was chosen as $\tau = \varrho^2/24 = 0.4167$.

	Location		Function value		PDE time step		Discrete Oja median	
	x_0	y_0	u	v	τu_t	τv_t	$u^* - u$	$v^* - v$
a)	0.9986	0.0523	0.9973	1.0000	0.002 495	0.000 417	0.001 896	0.000 538
b)	0.9659	0.2588	0.9330	1.0000	0.002 388	0.000 417	0.002 355	0.000 417
c)	0.7071	0.7071	0.5000	1.0000	0.001 667	0.000 417	0.001 650	0.000 404
d)	0.2588	0.9659	0.0670	1.0000	0.000 945	0.000 417	0.000 943	0.000 407
e)	0.0523	0.9986	0.0027	1.0000	0.000 838	0.000 417	0.000 920	0.000 448
f)	0.1000	0.1000	0.0100	0.1414	0.001 667	0.002 946	0.001 587	0.003 668
g)	0.3000	0.3000	0.0900	0.4243	0.001 667	0.000 982	0.001 654	0.001 009

References

1. L. Alvarez, P.-L. Lions, and J.-M. Morel. Image selective smoothing and edge detection by nonlinear diffusion. II. *SIAM Journal on Numerical Analysis*, 29:845–866, 1992.

2. J. Astola, P. Haavisto, and Y. Neuvo. Vector median filters. *Proceedings of the IEEE*, 78(4):678–689, 1990.

3. T. L. Austin. An approximation to the point of minimum aggregate distance. *Metron*, 19:10–21, 1959.

4. B. Chakraborty and P. Chaudhuri. On a transformation and re-transformation technique for constructing an affine equivariant multivariate median. *Proceedings of the AMS*, 124(6):2539–2547, 1996.

5. B. Chakraborty and P. Chaudhuri. A note on the robustness of multivariate medians. *Statistics and Probability Letters*, 45:269–276, 1999.

6. D. H. Chung and G. Sapiro. On the level lines and geometry of vector-valued images. *IEEE Signal Processing Letters*, 7(9):241–243, 2000.

7. U. Eckhardt. Root images of median filters. *Journal of Mathematical Imaging and Vision*, 19:63–70, 2003.

8. C. Gini and L. Galvani. Di talune estensioni dei concetti di media ai caratteri qualitativi. *Metron*, 8:3–209, 1929.

9. F. Guichard and J.-M. Morel. Partial differential equations and image iterative filtering. In I. S. Duff and G. A. Watson, editors, *The State of the Art in Numerical Analysis*, number 63 in IMA Conference Series (New Series), pages 525–562. Clarendon Press, Oxford, 1997.

10. J. F. Hayford. What is the center of an area, or the center of a population? *Journal of the American Statistical Association*, 8(58):47–58, 1902.

11. T. P. Hettmansperger and R. H. Randles. A practical affine equivariant multivariate median. *Biometrika*, 89(4):851–860, 2002.

12. A. Kleefeld, M. Breuß, M. Welk, and B. Burgeth. Adaptive filters for color images: median filtering and its extensions. In *Proc. 5th Computational Color Imaging Workshop (CCIW '15)*, Saint Etienne, France, Mar. 2015. In press.

13. H. Oja. Descriptive statistics for multivariate distributions. *Statistics and Probability Letters*, 1:327–332, 1983.

14. C. R. Rao. Methodology based on the l_1-norm in statistical inference. *Sankhyā A*, 50:289–313, 1988.

15. C. G. Small. A survey of multidimensional medians. *International Statistical Review*, 58(3):263–277, 1990.

16. C. Spence and C. Fancourt. An iterative method for vector median filtering. In *Proc. 2007 IEEE International Conference on Image Processing*, volume 5, pages 265–268, 2007.

17. J. W. Tukey. *Exploratory Data Analysis*. Addison–Wesley, Menlo Park, 1971.

18. A. Weber. *Über den Standort der Industrien*. Mohr, Tübingen, 1909.

19. E. Weiszfeld. Sur le point pour lequel la somme des distances de n points donnés est minimum. *Tôhoku Mathematics Journal*, 43:355–386, 1937.

20. M. Welk and M. Breuß. Morphological amoebas and partial differential equations. In P. W. Hawkes, editor, *Advances in Imaging and Electron Physics*, volume 185, pages 139–212. Elsevier Academic Press, 2014.

21. M. Welk, M. Breuß, and O. Vogel. Morphological amoebas are self-snakes. *Journal of Mathematical Imaging and Vision*, 39:87–99, 2011.

22. M. Welk, C. Feddern, B. Burgeth, and J. Weickert. Median filtering of tensor-valued images. In B. Michaelis and G. Krell, editors, *Pattern Recognition*, volume 2781 of *Lecture Notes in Computer Science*, pages 17–24. Springer, Berlin, 2003.