Supplementary material to the paper "Template Matching via Densities on the Roto-Translation Group"

Erik J. Bekkers, Marco Loog, Bart M. ter Haar Romeny, and Remco Duits

1 Probabilistic Interpretation of the Smoothing Prior in $SE(2)$

In this section we relate the $SE(2)$ smoothing prior to time resolvent hypo-elliptic[1] diffusion processes on $SE(2)$. First we aim to familiarize the reader with the concept of resolvent diffusions on $\mathbb{R}^2$ in Subsec. 1.1. Then we pose in Subsec. 1.2 a new problem (the single patch problem), which is a special case of our $SE(2)$ linear regression problem, that we use to link the left-invariant regularizer to tue resolvents of hypo-elliptic diffusions on $SE(2)$.

1.1 Resolvent Diffusion Processes

A classic approach to noise suppression in images is via diffusion regularizations with PDE's of the form [3]

$$\left\{ \begin{array}{ll} \frac{\partial}{\partial \tau} u & = \Delta u, \\ u|_{\tau=0} & = u_0, \end{array} \right. \tag{1}$$

where Δ denotes the Laplace operator. Solving (1) for any diffusion time $\tau > 0$ gives a smoothed version of the input u_0. The time-resolvent process of the PDE is defined by the Laplace transform with respect to τ; time τ is integrated out using a memoryless negative exponential distribution $P(\mathcal{T} = \tau) = \alpha e^{-\alpha \tau}$. Then, the time integrated solutions

$$t(\mathbf{x}) = \alpha \int_0^\infty u(\mathbf{x}, \tau) e^{-\alpha \tau} \mathrm{d}\tau,$$

with decay parameter α, are in fact the solutions

$$t = \underset{t \in \mathbb{L}_2(\mathbb{R}^2)}{\operatorname{argmin}} \left[\|t - t_0\|^2_{\mathbb{L}_2(\mathbb{R}^2)} + \lambda \int_{\mathbb{R}^2} \|\nabla t(\tilde{\mathbf{x}})\|^2 \, \mathrm{d}\tilde{\mathbf{x}} \right], \tag{2}$$

with $\lambda = \alpha^{-1}$, and corresponding Euler-Lagrange equation

$$(I - \lambda \Delta) t = t_0 \quad \Leftrightarrow \quad t = \lambda^{-1} \left(\frac{1}{\lambda} - \Delta \right)^{-1} t_0, \tag{3}$$

to which we refer as the "resolvent" equation [4], as it involves operator $(\alpha I - \Delta)^{-1}$, $\alpha = \lambda^{-1}$. In the next subsections, we follow a similar procedure with $SE(2)$ instead of $\mathbb{R}^2$, and show how the smoothing regularizer in Eq. (28) and (30) of the main article relates to Laplace transforms of hypo-elliptic diffusions on the group $SE(2)$ [2], [5].

1.2 The Fundamental Single Patch Problem

In order to grasp what the (anisotropic regularization term) in Eq. (28) and (30) of the main article actually means in terms of stochastic interpretation/probabilistic line propagation, let us consider the following single patch problem and optimize

$$\begin{aligned} \mathcal{E}_{sp}(T) = & \left| \left(G_s *_{\mathbb{R}^2} T(\cdot, \cdot, \theta_0) \right)(\mathbf{x}_0) - 1 \right|^2 \\ & + \lambda \int_{\mathbb{R}^2} \int_0^{2\pi} \|\nabla T(\tilde{\mathbf{x}}, \tilde{\theta})\|^2_D \mathrm{d}\tilde{\mathbf{x}} \mathrm{d}\tilde{\theta} + \mu \|T\|^2_{\mathbb{L}_2(SE(2))}, \end{aligned} \tag{4}$$

with $(\mathbf{x}_0, \theta_0) = g_0 := (x_0, y_0, \theta_0) \in SE(2)$ the fixed center of the template, and with spatial Gaussian kernel

$$G_s(\mathbf{x}) = \frac{1}{4\pi s} e^{-\frac{\|\mathbf{x}\|^2}{4s}}.$$

Regarding this problem, we note the following:

- In the original problem (28) of the main article we take $N = 1$, with
$$U_{f_1}(x, y, \theta) = G_s(x - x_0, y - y_0) \, \delta_{\theta_0}(\theta) \tag{5}$$
representing a local spatially smoothed spike in $SE(2)$, and set $y_1 = 1$. The general single patch case (for arbitrary U_{f_1}) can be deduced by superposition of such impulse responses.
- We use $\mu > 0$ to suppress the output elsewhere.
- We use $0 < s \ll 1$. This minimum scale due to sampling removes the singularity at $(\mathbf{0}, 0)$ from the kernel that solves (4), as proven in [2].

Theorem 1. *The solution to the single patch problem (4) coincides up to scalar multiplication with the time integrated hypo-elliptic Brownian motion kernel on $SE(2)$ (depicted in Fig. 1).*

- *E.J. Bekkers and B.M. ter Haar Romeny are with the department of Biomedical Engineering, Eindhoven University of Technology (TU/e), the Netherlands. E-mail: {e.j.bekkers,b.m.terhaarromeny}@tue.nl*
- *M. Loog is with the Pattern Recognition Laboratory, Delft University of Technology, the Netherlands. E-mail: m.loog@tudelft.nl*
- *B.M. ter Haar Romeny is also with the department of Biomedical and Information Engineering, Northeastern University, Shenyang, China.*
- *R. Duits is with the department of Mathematics and Computer Science, TU/e; he is also affiliated to the department of Biomedical Enginering, TU/e. Email: r.duits@tue.nl*

Manuscript received ...., ...; revised ...., ...

1. This diffusion process on $SE(2)$ is called *hypo-elliptic* as its generator equals $(\partial_\xi)^2 + D_{\theta\theta}(\partial_\theta)^2$ and diffuses only in 2 directions in a 3D space. This boils down to a sub-Riemannian manifold structure [1], [2]. Smoothing in the missing (∂_η) direction is achieved via the commutator: $[\partial_\theta, \cos\theta \partial_x + \sin\theta \partial_y] = -\sin\theta \partial_x + \cos\theta \partial_y$.

Proof. We optimize $\mathcal{E}_{sp}(T)$ over the set $\mathcal{S}(SE(2))$ of all functions $T : SE(2) \to \mathbb{R}$ that are bounded and on $SE(2)$, infinitely differentiable on $SE(2)\setminus\{g_0\}$, and rapidly decreasing in spatial direction, and 2π periodic in θ. We omit topological details on function spaces and Hörmanders condition [6]. Instead, we directly proceed with applying the Euler-Lagrange technique to the single patch problem:

$$\forall_{\delta\in\mathcal{S}(SE(2))} : \lim_{\epsilon\downarrow 0}\left\{\frac{\mathcal{E}_{sp}(T+\epsilon\delta)-\mathcal{E}_{sp}(T)}{\epsilon}\right\} = 0 \Leftrightarrow$$
$$(S_s^*S_s + \lambda R + \mu I)T = S_s^* y_1 = S_s^* 1, \quad (6)$$

with linear functional (distribution) $\mathcal{S}_s$ given by

$$(S_s T) = (G_s *_{\mathbb{R}^2} T(\cdot,\theta_0))(\mathbf{x}_0),$$

and with regularization operator R given by

$$R = -\Delta_{SE(2)} := -(D_{\theta\theta}\partial_\theta^2 + D_{\xi\xi}\partial_\xi^2 + D_{\eta\eta}\partial_\eta^2) \geq 0.$$

Note that $\lim_{s\to 0} S_s = \delta_{(\mathbf{x}_0,\theta_0)}$ in distributional sense, and that the constraint $s > 0$ is crucial for solutions T to be bounded at $(\mathbf{x}_0,\theta_0)$. By definition the adjoint operator S_s^* is given by

$$\begin{aligned}(S_s^* y, T)_{\mathbb{L}_2(SE(2))} &= (y, S_s T) = y \int_{\mathbb{R}^2} G_s(\mathbf{x}-\mathbf{x}_0)T(\mathbf{x},\theta_0)\,\mathrm{d}\mathbf{x}\\ &= y \int_0^{2\pi}\int_{\mathbb{R}^2} G_s(\mathbf{x}-\mathbf{x}_0)\delta_{\theta_0}(\theta)T(\mathbf{x},\theta)\,\mathrm{d}\mathbf{x}\mathrm{d}\theta,\\ &= (y\, G_s(\cdot-\mathbf{x}_0)\delta_{\theta_0}(\cdot), T)_{\mathbb{L}_2(SE(2))}\end{aligned}$$

and thereby we deduce that

$$(S_s^* y)(\mathbf{x},\theta) = y\, G_s(\mathbf{x}-\mathbf{x}_0)\delta_{\theta_0}(\theta),$$
$$S_s^*(S_s T) = T_0^s\, G_s(\mathbf{x}-\mathbf{x}_0)\delta_{\theta_0}(\theta),$$

with $\infty > T_0^s := (G_s *_{\mathbb{R}^2} T(\cdot,\theta_0))(\mathbf{x}_0) > 1$ for $0 < s \ll 1$. The Euler-Lagrange equation (6) becomes

$$(-\lambda\Delta_{SE(2)} + \mu I)T = (1 - T_0^s)G_s(\mathbf{x}-\mathbf{x}_0)\delta_{\theta_0}(\theta).$$

Now, when setting $T_{new} = \frac{T}{1-T_0^s}$ we arrive at the hypo-elliptic resolvent equation on $SE(2)$:

$$\begin{aligned}(-\lambda\Delta_{SE(2)} + \mu I)T_{new} = (G_s *_{\mathbb{R}^2} \delta_{\mathbf{x}_0})\delta_{\theta_0} \quad &\Leftrightarrow\\ T_{new} = \left(-\lambda\Delta_{SE(2)} + \mu I\right)^{-1} e^{s\Delta_{\mathbb{R}^2}}\delta_{g_0} \quad &(7)\\ = e^{s\Delta_{\mathbb{R}^2}}\left(-\lambda\Delta_{SE(2)} + \mu I\right)^{-1}\delta_{g_0}\end{aligned}$$

where we write $e^{s\Delta_{\mathbb{R}^2}} f = G_s *_{\mathbb{R}^2} f$ for the diffusion operator, to stress the vanishing commutators

$$[e^{s\Delta_{\mathbb{R}^2}}, \Delta_{SE(2)}] = e^{s\Delta_{\mathbb{R}^2}}\Delta_{SE(2)} - \Delta_{SE(2)}e^{s\Delta_{\mathbb{R}^2}} = 0,$$

which directly follows from $[\Delta_{\mathbb{R}^2}, \Delta_{SE(2)}] = 0$. In fact, from these vanishing commutators one can deduce that, thanks to the isotropy of Gaussian kernel, blurring with inner-scale $s > 0$ can be done either before applying the resolvent operator or after (as seen in (7)).

The solutions T_{new} are precisely the probabilistic kernels $R_{\alpha,s} : SE(2) \to \mathbb{R}$ for time integrated contour enhancements studied in [2], [5]. In fact we see that

$$T_{new}(g) = \mu^{-1} R_{\alpha,s}(g_0^{-1}g),$$

where $R_{\alpha,s} = (I - \alpha^{-1}\Delta_{SE(2)})^{-1}e^{s\Delta_{\mathbb{R}^2}}\delta_{(\mathbf{0},0)}$ (i.e., the impuls response of the resolvent operator) denotes the time-integration of the hypo-elliptic diffusion kernel $K_{\tau,s} = e^{\tau\Delta_{SE(2)}}e^{s\Delta_{\mathbb{R}^2}}\delta_{(\mathbf{0},0)}$:

$$R_{\alpha,s}(g) = \alpha\int_0^\infty K_{\tau,s}(g)\, e^{-\alpha\tau}\,\mathrm{d}\tau,$$

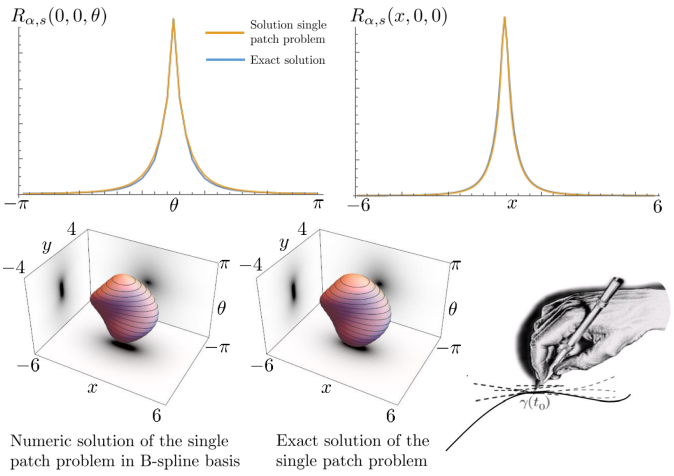

Fig. 1. Top row: Comparison of kernel $R_{\alpha,s}(x,y,\theta)$ along respectively the θ and x axis. Bottom row: Isosurface of the kernel computed by solving the fundamental single patch problem (4), the exact solution, and an illustration of the drunkman's pencil. For Monte Carlo simulations of the drunkman's pencil see the supplementary materials.

for which 3 different exact analytic formulas are derived in [5]. The kernel $R_{\alpha,s}(\mathbf{x},\theta)$ denotes the probability density of finding a random brush stroke (regardless its traveling time) at location $\mathbf{x}$ with orientation θ given that a 'drunkman's pencil' starts at $g = (\mathbf{0},0)$ at time zero. Here the traveling time τ of the random pencil is assumed to be negatively exponentially distributed with expectation α^{-1}. □

1.3 Expansion in B-splines

Now we consider the B-spline expansions (Eq. (34) in the main article) and apply our optimization algorithm (cf. Subsec. 2.4 of the main article) to the single patch problem (4), with $(\mathbf{x}_0,\theta_0) = (\mathbf{0},0)$. Here we no longer need a smoothing with a continuous Gaussian G_s, as expansion in the B-spline basis already includes regularization. Now we set for the smooth spike $U_{f_1}(x,y,\theta) = B^n\left(\frac{x}{s_k}\right)B^n\left(\frac{y}{s_l}\right)B^n\left(\frac{\theta \bmod 2\pi}{s_m}\right)$, and we thus approximate spikes by the same B-spline basis in which we expressed our templates. We accept extra regularization (like we did with the Gaussian in the previous section) and choose to represent a spike by a normal B-spline. After all, via the central limit theorem B-splines converge to Gaussians when increasing n. We also considered to instead use the fundamental B-Spline [7, Fig. 2], which is sharper but suffers from oscillations, yielding less favorable results.

In our normal B-spline setting, this choice of smooth spike representation (cf. Eq. (14) in the main article) leads to the following equations

$$(S^\dagger S + \lambda R + \mu I)T = S^\dagger 1,$$

with S the $[1 \times N_k N_l N_m]$-matrix whose components are given by $M(0,0,0)\,B_{s_k s_l s_m}(k,l,m)$. Akin to the previous derivations (7) this matrix-equation can be rewritten as

$$(\lambda R + \mu I)\, T_{new} = S^\dagger 1.$$

In particular our B-spline basis algorithm is a new algorithm that can be used for the resolvent (hypo-)elliptic diffusion process on $SE(2)$. The benefit over Fourier based

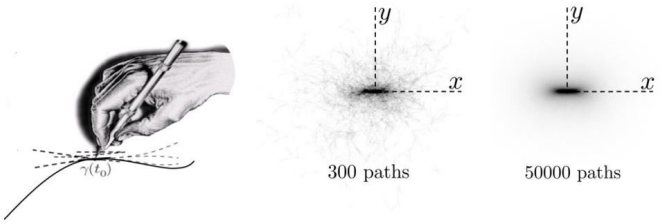

Fig. 2. Stochastic random process for contour enhancement.

algorithms is the local support of the basis functions, which allow for sparse representations.

In Fig. 1 we compare the impulse response for Tikhonov regularization via our B-spline expansion algorithm with the Brownian motion prior on $SE(2)$ (using a fine B-spline basis) to the exact solutions derived in [2], [5]. The strong accuracy of our algorithm shows that even in the discrete B-spline setting the probabilistic interpretation (Thm. 1) of our prior in $SE(2)$-template matching holds.

1.4 The Drunkman's Pencil

Similar to the diffusions on $\mathbb{R}^2$, given by (1), the hypo-elliptic diffusion process on $SE(2)$ is described by the following PDE:

$$\begin{cases} \frac{\partial}{\partial \tau} W &= (D_{\xi\xi}\partial_\xi^2 + D_{\theta\theta}\partial_\theta^2)W, \\ W|_{\tau=0} &= W_0, \end{cases} \tag{8}$$

initialized with $W_0 \in \mathbb{L}_2(\mathbb{R}^2)$ at time $\tau = 0$. The PDE can be used to obtain the solutions of our single patch problem by initializing W_0 with a smooth spike such as we did in Subsec. 1.3, e.g. taking $W_0 = U_{f_1}(x,y,\theta) = B^n\left(\frac{x}{s_k}\right) B^n\left(\frac{y}{s_l}\right) B^n\left(\frac{\theta \bmod 2\pi}{s_m}\right)$.

The PDE in (8) is the forward Kolmogorov equation [8] of the following stochastic process [2]:

$$\begin{cases} \mathbf{x}(\tau) = \mathbf{x}(0) + \\ \qquad \sqrt{2D_{\xi\xi}}\,\epsilon_\xi \int_0^\tau (\cos\theta(\tau)\mathbf{e}_x + \sin\theta(\tau)\mathbf{e}_y)\frac{1}{2\sqrt{\tau}}\mathrm{d}\tau \\ \theta(\tau) = \theta(0) + \sqrt{\tau}\sqrt{2D_{\theta\theta}}\,\epsilon_\theta, \qquad \epsilon_\xi, \epsilon_\theta\, \mathcal{N}(0,1), \end{cases} \tag{9}$$

where ϵ_ξ and ϵ_θ are sampled from a normal distribution with expectation 0 and unit standard deviation. The stochastic process in (9) can be interpreted as the motion of a drunkman's pencil: it randomly moves forward and backwards, and randomly changes its orientation along the way. The resolvent hypo-elliptic diffusion kernels $R_{\alpha,s}(g)$ (solutions to the fundamental single patch problem, up to scalar multiplication) can then also be obtained via Monte Carlo simulations, where the stochastic process is sampled many times with a negatively exponentially distributed traveling time ($P(\mathcal{T} = \tau) = \alpha e^{-\alpha\tau}$) in order to be able to estimate the probability density kernel $R_{\alpha,s}(g)$. This process is illustrated in Fig. 2.

2 The Smoothing Regularization Matrix R

When expanding the templates t and T in a finite B-Spline basis (Sec. 2 and 3 of the main article), the energy functionals (7), (11), (28) and (30) of the main article can be expressed in matrix vector form. The following theorems summarize how to compute the matrix R, which encodes the smoothing prior, for respectively the $\mathbb{R}^2$ and $SE(2)$ case.

Lemma 1. *The discrete smoothing regularization-term of energy functional (7) of the main article can be expressed directly in the B-Spline coefficients* $\mathbf{c}$ *as follows*

$$\iint_{\mathbb{R}^2} \|\nabla t(x,y)\|^2 \mathrm{d}x\mathrm{d}y = \mathbf{c}^\dagger R \mathbf{c}, \tag{10}$$

with $\mathbf{c}$ *given by Eq. (16) of the main article, and with*

$$R = R_x^k \otimes R_x^l + R_y^k \otimes R_y^l, \tag{11}$$

a $[N_k N_l \times N_k N_l]$ *matrix. The elements of the matrices in (11) are given by*

$$\begin{array}{rl} R_x^k(k,k') &= -\frac{1}{s_k}\frac{\partial^2 B^{2n+1}}{\partial x^2}(k'-k) \\ R_x^l(l,l') &= s_l B^{2n+1}(l'-l), \\ R_y^k(k,k') &= s_k B^{2n+1}(k'-k), \\ R_y^l(l,l') &= -\frac{1}{s_l}\frac{\partial^2 B^{2n+1}}{\partial y^2}(l'-l). \end{array} \tag{12}$$

Proof. For the sake of readability we divide the regularization-term in two parts:

$$\begin{aligned} \iint_{\mathbb{R}^2} \|\nabla t(x,y)\|^2 \mathrm{d}x\mathrm{d}y &= \iint_{\mathbb{R}^2} \left|\tfrac{\partial t}{\partial x}(x,y)\right|^2 \\ &\quad + \left|\tfrac{\partial t}{\partial y}(x,y)\right|^2 \mathrm{d}x\mathrm{d}y, \end{aligned} \tag{13}$$

where

$$\begin{array}{rl} \mathcal{R}_x &= \iint_{\mathbb{R}^2} \left|\tfrac{\partial t}{\partial x}(x,y)\right|^2 \mathrm{d}x\mathrm{d}y, \text{ and} \\ \mathcal{R}_y &= \iint_{\mathbb{R}^2} \left|\tfrac{\partial t}{\partial y}(x,y)\right|^2 \mathrm{d}x\mathrm{d}y. \end{array}$$

We first derive the matrix-vector representation of $\mathcal{R}_x$ as follows:

$$\begin{aligned} \mathcal{R}_x &= \iint_{\mathbb{R}^2} \left|\tfrac{\partial t}{\partial x}(x,y)\right|^2 \mathrm{d}x\mathrm{d}y \\ &= \sum_{k,k'=1}^{N_k} \sum_{l,l'=1}^{N_l} \iint_{\mathbb{R}^2} \overline{c_{k,l}\tfrac{\partial B^n}{\partial x}(\tfrac{x}{s_k} - k)B^n(\tfrac{y}{s_l} - l)} \\ &\qquad c_{k,l}\tfrac{\partial B^n}{\partial x}(\tfrac{x}{s_k} - k')B^n(\tfrac{y}{s_l} - l')\mathrm{d}x\mathrm{d}y \\ &= \sum_{k,k'=1}^{N_k} \sum_{l,l'=1}^{N_l} \overline{c_{k,l}} c_{k',l'} \\ &\qquad \left[\int_{-\infty}^{\infty} \tfrac{\partial B^n}{\partial x}(\tfrac{x}{s_k} - k)\tfrac{\partial B^n}{\partial x}(\tfrac{x}{s_k} - k')\mathrm{d}x\right] \\ &\qquad \left[\int_{-\infty}^{\infty} B^n(\tfrac{y}{s_l} - l)B^n(\tfrac{y}{s_l} - l')\mathrm{d}y\right] \\ &\overset{1}{=} \sum_{k,k'=1}^{N_k} \sum_{l,l'=1}^{N_l} \overline{c_{k,l}} c_{k',l'} \left[\tfrac{1}{s_k}\left(\tfrac{\partial B^n}{\partial x} * \tfrac{\partial B^n}{\partial x}\right)(k'-k)\right] \\ &\qquad \left[s_l\left(B^n * B^n\right)(l'-l)\right] \\ &\overset{2}{=} \sum_{k,k'=1}^{M} \sum_{l,l'=1}^{N} \overline{c_{k,l}} c_{k',l'} \left[\tfrac{1}{s_k}\tfrac{\partial^2 B^{2n+1}}{\partial x^2}(k'-k)\right] \\ &\qquad \left[s_l B^{2n+1}(l'-l)\right] \\ &= \mathbf{c}^\dagger (R_x^k \otimes R_x^l)\mathbf{c}. \end{aligned} \tag{14}$$

Here the following properties are used:

1) The integrals of shifted B-splines can be expressed as convolutions:

$$\begin{aligned} &\int_{-\infty}^{\infty} \frac{\partial B^n}{\partial x}\left(\frac{x}{s_k} - k\right)\frac{\partial B^n}{\partial x}\left(\frac{x}{s_k} - k'\right)\mathrm{d}x \\ &\quad = -\frac{1}{s_k}\int_{-\infty}^{\infty} \frac{\partial B^n}{\partial u}(u)\frac{\partial B^n}{\partial u}((k'-k) - u)\mathrm{d}u \\ &\quad = -\frac{1}{s_k}\left(\frac{\partial B^n}{\partial u} * \frac{\partial B^n}{\partial u}\right)(k'-k). \end{aligned}$$

This is easily verified by substitution of integration variable ($u = -\frac{x}{s_k} + k$) and noting that $B^n(x) = B^n(-x)$ and $\frac{\partial B^n}{\partial x}(x) = -\frac{\partial B^n}{\partial x}(-x)$.

2) Convolution of two B-splines B^n of order n results in a B-Spline B^{2n+1} of order $2n+1$:

$$B^n * B^n = B^{2n+1}.$$

The elements of the matrices R_y^k and R_y^l are derived in a similar manner. □

As a result of Lemma 1 we can state the following.

Corollary 1. *Let $V = \text{span}\{B^n_{k,l}\}$, with $k = 1, \ldots, N_k, l = 1, \ldots, N_l$, and shifted B-splines $B^n_{k,l}$ of order n (cf. Subsec. 2.4 of the main article). Let the energy function $E^B_{lin} : \mathbb{R}^{N_k N_l} \to \mathbb{R}^+$ be given by Eq. (13) of the main article. Then the optimal continuous template of the constrained optimization problem (cf. Eq. (7) of the main article)*

$$t^* = \underset{t \in V}{\text{argmin}} \;\; E_{lin}(t)$$

has coefficients $\mathbf{c}^*$ w.r.t. the B-spline basis for V, that are the unique solution of

$$\nabla_{\mathbf{c}} E^B(\mathbf{c}^*) = \mathbf{0},$$

which boils down to Eq. (14) of the main article.

Lemma 2. *The discrete regularization-term of energy functional (28) of the main article can be expressed directly in the B-Spline coefficients:*

$$\iiint_{SE(2)} \|\nabla T\|_D \, \mathrm{d}x\mathrm{d}y\mathrm{d}\theta = \mathbf{c}^T (D_{\xi\xi} R_\xi + D_{\eta\eta} R_\eta + D_{\theta\theta} R_\theta)\mathbf{c}. \tag{15}$$

Matrix R_ξ is given by

$$\begin{aligned} R_\xi = &\left(R_\xi^{Ix} \otimes R_\xi^{Iy} \otimes R_\xi^{I\theta}\right) + \left(R_\xi^{IIx} \otimes R_\xi^{IIy} \otimes R_\xi^{II\theta}\right) \\ &+ \left(R_\xi^{IIIx} \otimes R_\xi^{IIIy} \otimes R_\xi^{III\theta}\right) + \left(R_\xi^{IVx} \otimes R_\xi^{IVy} \otimes R_\xi^{IV\theta}\right) \end{aligned} \tag{16}$$

with the elements of the matrices used in the Kronecker products given by

$$\begin{aligned} R_\xi^{Ix}(k,k') &= -\tfrac{1}{s_k} \tfrac{\partial^2 B^{2n+1}}{\partial x^2}(k'-k), \\ R_\xi^{Iy}(l,l') &= s_l B^{2n+1}(l'-l), \\ R_\xi^{I\theta}(m,m') &= \int_0^\pi \cos^2(\theta) B^n(\tfrac{\theta}{s_m} - m) B^n(\tfrac{\theta}{s_m} - m') \mathrm{d}\theta, \end{aligned} \tag{17}$$

$$\begin{aligned} R_\xi^{IIx}(k,k') &= -R_\xi^{IIIx}(k,k') = \tfrac{\partial B^{2n+1}}{\partial x}(k'-k), \\ R_\xi^{IIy}(l,l') &= -R_\xi^{IIIy}(l,l') = -\tfrac{\partial B^{2n+1}}{\partial y}(l'-l), \\ R_\xi^{II\theta}(m,m') &= R_\xi^{III\theta}(m,m') = \\ &\int_0^\pi \cos(\theta)\sin(\theta) B^n(\tfrac{\theta}{s_m} - m) B^n(\tfrac{\theta}{s_m} - m') \mathrm{d}\theta, \end{aligned} \tag{18}$$

$$\begin{aligned} R_\xi^{IVx}(k,k') &= s_k B^{2n+1}(k'-k), \\ R_\xi^{IVy}(l,l') &= -\tfrac{1}{s_l} \tfrac{\partial^2 B^{2n+1}}{\partial y^2}(l'-l), \\ R_\xi^{IV\theta}(m,m') &= \int_0^\pi \sin^2(\theta) B^n(\tfrac{\theta}{s_m} - m) B^n(\tfrac{\theta}{s_m} - m') \mathrm{d}\theta. \end{aligned} \tag{19}$$

Matrix R_η is given by

$$\begin{aligned} R_\eta = &\left(R_\xi^{IIx} \otimes R_\xi^{IIy} \otimes R_\xi^{IV\theta}\right) - \left(R_\xi^{IIx} \otimes R_\xi^{IIy} \otimes R_\xi^{II\theta}\right) \\ &- \left(R_\xi^{IIIx} \otimes R_\xi^{IIIy} \otimes R_\xi^{III\theta}\right) + \left(R_\xi^{IVx} \otimes R_\xi^{IVy} \otimes R_\xi^{I\theta}\right). \end{aligned} \tag{20}$$

Matrix R_θ is given by

$$R_\theta = \left(R_\theta^x \otimes R_\theta^y \otimes R_\theta^\theta\right), \tag{21}$$

with the elements of the matrices given by

$$\begin{aligned} R_\theta^x(k,k') &= s_k B^{2n+1}(k'-k), \\ R_\theta^y(l,l') &= s_l B^{2n+1}(l'-l), \\ R_\theta^\theta(m,m') &= -\tfrac{1}{s_m} \tfrac{\partial^2 B^{2n+1}}{\partial \theta^2}(m'-m). \end{aligned} \tag{22}$$

Proof. The proof of Lemma 2 follows the same steps as in the proof of Lemma 1, only here left-invariant derivatives are used (cf. Eq. (24) of the main article). The four separate terms $I - IV$ of Eq. (16) arise from the left invariant derivative ∂_ξ: $\left|\frac{\partial T}{\partial \xi}\right|^2 = \left|\cos(\theta)\frac{\partial T}{\partial x} + \sin(\theta)\frac{\partial T}{\partial y}\right|^2$. □

Lemma 2 has the following consequence.

Corollary 2. *Let $V = \text{span}\{B^n_{k,l,m}\}$, with $k = 1, \ldots, N_k, l = 1, \ldots, N_l, m = 1, \ldots, N_m$, and shifted B-splines $B^n_{k,l,m}$ of order n (cf. Subsec. 3.5 of the main article). Let the energy function $\mathcal{E}^B_{lin} : \mathbb{R}^{N_k N_l N_m} \to \mathbb{R}^+$ be given by*

$$\mathcal{E}^B_{lin}(\mathbf{c}) = \frac{1}{N}\|S\mathbf{c} - \mathbf{y}\|^2 + \mathbf{c}^\dagger(\lambda R + \mu I)\mathbf{c}$$

With S and $\mathbf{y}$ given by (33) of the main article and with $R = D_{\xi\xi}R_\xi + D_{\eta\eta}R_\eta + D_{\theta\theta}R_\theta$ given in Lemma 2. Then the optimal continuous template of the constrained optimization problem (cf. Eq. (28) of the main article)

$$T^* = \underset{T \in V}{\text{argmin}} \;\; \mathcal{E}_{lin}(T)$$

has coefficients $\mathbf{c}^*$ w.r.t. the B-spline basis for V that are the unique solution of

$$\nabla_{\mathbf{c}} \mathcal{E}^B(\mathbf{c}^*) = \mathbf{0}.$$

This boils down to Eq. (14) of the main article, but then on $\mathbb{R}^{N_k N_l N_m}$ with matrices R and S given above.

3 Normalized Cross Correlation

In most applications it is necessary to make the detection system invariant to local contrast and luminosity changes. In our template matching framework this can either be achieved via certain pre-processing steps that filter out these variations, or by means of normalized cross-correlation. In normalized cross-correlation, both the template as well as the image are (locally) normalized to zero mean and unit standard deviation (with respect to the inner product used in the cross-correlations). In this section, we explain the necessary adaptations to extend the standard cross-correlation based framework to normalized cross-correlations.

3.1 Normalized Cross-Correlation in $\mathbb{R}^2$

In the usual cross-correlation based template matching approach, as described in Sec. 2 of the main article, we rely on the standard $\mathbb{L}_2(\mathbb{R}^2)$ inner product (Eq. (6) of the main article). In normalized cross-correlation it is however convenient to extend this inner product to include a windowing function m which indicates the relevant region (support) of the template. As such, the inner product with respect to windowing function m is given by

$$(t, f)_{\mathbb{L}_2(\mathbb{R}^2, m d\tilde{\mathbf{x}})} := \int_{\mathbb{R}^2} \overline{t(\tilde{\mathbf{x}})} f(\tilde{\mathbf{x}}) m(\tilde{\mathbf{x}}) \mathrm{d}\tilde{\mathbf{x}}, \tag{23}$$

with associated norm $\|\cdot\|_{\mathbb{L}_2(\mathbb{R}^2, m d\tilde{\mathbf{x}})} = \sqrt{(\cdot,\cdot)_{\mathbb{L}_2(\mathbb{R}^2, m d\tilde{\mathbf{x}})}}$. The windowing function has to be a smooth function $m : \mathbb{R}^2 \to \mathbb{R}^+$ with $\int_{\mathbb{R}^2} m(\tilde{\mathbf{x}}) \mathrm{d}\tilde{\mathbf{x}} = 1$. In this work, the use of a window m is also convenient to deal with boundary conditions in the optimization problems for template construction. We define

$$m(\mathbf{x}) := \varsigma \, e^{-\frac{\|\mathbf{x}\|^2}{s}} \sum_{i=0}^{n} \frac{(\|\mathbf{x}\|^2/s)^i}{i!}, \tag{24}$$

which smoothly approximates the indicator function $1_{[0,r]}(\|\mathbf{x}\|)$, covering a disk with radius r, when setting $s = \frac{2r^2}{1+2n}$, see e.g. [9, Fig. 2]. The constant ς normalizes the function such that $\int_{\mathbb{R}^2} m(\tilde{\mathbf{x}}) \mathrm{d}\tilde{\mathbf{x}} = 1$.

In normalized cross-correlation the image is locally normalized (at position $\mathbf{x}$) to zero mean and unit standard deviation, which is done as follows

$$\hat{f}_{\mathbf{x}}(\tilde{\mathbf{x}}) := \frac{f(\tilde{\mathbf{x}}) - \langle f \rangle_{\mathcal{T}_{\mathbf{x}} m}}{\|f(\tilde{\mathbf{x}}) - \langle f \rangle_{\mathcal{T}_{\mathbf{x}} m}\|_{\mathbb{L}_2(\mathbb{R}^2, \mathcal{T}_{\mathbf{x}} m d\tilde{\mathbf{x}})}}, \tag{25}$$

with local mean $\langle f \rangle_m = (1, f)_{\mathbb{L}_2(\mathbb{R}^2, m d\tilde{\mathbf{x}})}$. Template $\hat{t}$ can be obtained via normalization of a given template t via

$$\hat{t}(\tilde{\mathbf{x}}) := \frac{t(\tilde{\mathbf{x}}) - \langle t \rangle_m}{\|t(\tilde{\mathbf{x}}) - \langle t \rangle_m\|_{\mathbb{L}_2(\mathbb{R}^2, m d\tilde{\mathbf{x}})}}. \tag{26}$$

Template matching is then done in the usual way (via (4) of the main article), however now $\hat{t}$ and $\hat{f}_{\mathbf{x}}$ are used instead of t and f. In fact, the entire $\mathbb{R}^2$ cross-correlation template matching, and template optimization framework is extended to normalized cross-correlation by substituting all instances of t with $\hat{t}$, f with $\hat{f}_{\mathbf{x}}$, and $(\cdot,\cdot)_{\mathbb{L}_2(\mathbb{R}^2)}$ with $(\cdot,\cdot)_{\mathbb{L}_2(\mathbb{R}^2, m d\tilde{\mathbf{x}})}$ in Sec. 2 of the main article. However, since templates $\hat{t}$ are directly constructed via the minimization of energy functionals, we will not explictely normalize the templates, unless they are obtained by other methods. E.g., Eq. (26) is used in the main article to construct basic templates obtained by averaging positive object patches (Subsec. 4.1 of the main article).

3.2 Normalized Cross-Correlation in $SE(2)$

Similar to the $\mathbb{R}^2$ case, templates and orientation scores are locally normalized to zero mean and unit standard deviation, however, now with respect to the $\mathbb{L}_2(SE(2), M d\tilde{g})$ inner product, which is given by

$$(T, U_f)_{\mathbb{L}_2(SE(2), M \mathrm{d}\tilde{g})} := \int_{\mathbb{R}^2} \int_0^{2\pi} \overline{T(\tilde{\mathbf{x}}, \tilde{\theta})} U_f(\tilde{\mathbf{x}}, \tilde{\theta}) M(\tilde{\mathbf{x}}, \tilde{\theta}) \mathrm{d}\tilde{\mathbf{x}} \mathrm{d}\tilde{\theta}, \tag{27}$$

with norm $\|\cdot\|_{\mathbb{L}_2(SE(2), M \mathrm{d}\tilde{g})} = \sqrt{(\cdot,\cdot)_{\mathbb{L}_2(SE(2), M \mathrm{d}\tilde{g})}}$. Also here windowing function M indicates the support of the template, and has the property $\int_{\mathbb{R}^2} \int_0^{2\pi} M(\tilde{\mathbf{x}}, \tilde{\theta}) \mathrm{d}\tilde{\mathbf{x}} \mathrm{d}\tilde{\theta} = 1$. We define

$$M(\mathbf{x}, \theta) := \frac{1}{2\pi} m(\mathbf{x}), \tag{28}$$

independent of θ and with $m(\mathbf{x})$ given by (24).

The (locally at g) normalized orientation score and template T are then given by

$$\hat{U}_{f,g}(\tilde{\mathbf{x}}, \tilde{\theta}) := \frac{U_f(\tilde{\mathbf{x}}, \tilde{\theta}) - \langle U_f \rangle_{\mathcal{L}_g M}}{\|U_f(\tilde{\mathbf{x}}, \tilde{\theta}) - \langle U_f \rangle_{\mathcal{L}_g M}\|_{\mathbb{L}_2(SE(2), \mathcal{L}_g M d\tilde{g})}}, \tag{29}$$

$$\hat{T}(\tilde{\mathbf{x}}, \tilde{\theta}) := \frac{T(\tilde{\mathbf{x}}, \tilde{\theta}) - \langle T \rangle_M}{\|T(\tilde{\mathbf{x}}, \tilde{\theta}) - \langle T \rangle_M\|_{\mathbb{L}_2(SE(2), M \mathrm{d}\tilde{g})}}, \tag{30}$$

with mean $\langle U_f \rangle_M = (1, U_f)_{\mathbb{L}_2(SE(2), M \mathrm{d}\tilde{g}))}$.

3.3 Efficient local normalization of $\hat{f}_{\mathbf{x}}$ and $\hat{U}_{f,g}$.

Since the normalized image $\hat{f}_{\mathbf{x}}$ depends on the location $\mathbf{x}$ it needs to be calculated for every translation of the template, which makes normalized cross-correlation computationally expensive. Therefore, (25) can be approximated by assuming that the local average is approximately constant in the area covered by m. That is, assuming $\langle f \rangle_{\mathcal{T}_{\mathbf{x}} m}(\tilde{\mathbf{x}}) \approx \langle f \rangle_{\mathcal{T}_{\tilde{\mathbf{x}}} m}(\tilde{\mathbf{x}}) = (m \star f)(\tilde{\mathbf{x}})$ for $\|\tilde{\mathbf{x}} - \mathbf{x}\| < r$, with r the radius that determines the extent of m, (25) is approximated as follows:

$$\hat{f}_{\mathbf{x}}(\tilde{\mathbf{x}}) \approx \frac{f(\tilde{\mathbf{x}}) - (m \star f)(\tilde{\mathbf{x}})}{\sqrt{(m \star (f - (m \star f))^2)(\tilde{\mathbf{x}})}}. \tag{31}$$

Similarly, in the $SE(2)$-case (29) can be approximated via

$$\hat{U}_{f,g}(\tilde{\mathbf{x}}, \tilde{\theta}) \approx \frac{U_f(\tilde{\mathbf{x}}, \tilde{\theta}) - (M \star_{SE(2)} U_f)(\mathbf{x}, \tilde{\theta})}{\sqrt{(M \star_{SE(2)} (U_f - (M \star_{SE(2)} U_f))^2)(\mathbf{x}, \tilde{\theta})}}. \tag{32}$$

3.4 Including a Region of Interest Mask

Depending on the application, large portions of the image might be masked out. This for example the case in retinal images (see circular masks in Fig. 4). To deal with this, template matching is only performed inside the region of interest defined by a mask image $m^{roi} : \mathbb{R}^2 \to \{0, 1\}$. Including such a mask is important in normalized template matching, and can be done by replacing the standard inner products by

$$(t, f)^{roi}_{\mathbb{L}_2(\mathbb{R}^2, m, d\tilde{\mathbf{x}})} := \frac{(t, f m^{roi})_{\mathbb{L}_2(\mathbb{R}^2, m, d\tilde{\mathbf{x}})}}{(1, m^{roi})_{\mathbb{L}_2(\mathbb{R}^2, m, d\tilde{\mathbf{x}})}}, \tag{33}$$

$$(T, U_f)^{roi}_{\mathbb{L}_2(SE(2), M, d\tilde{g})} := \frac{(T, U_f M^{roi})_{\mathbb{L}_2(SE(2), M, d\tilde{g})}}{(1, M^{roi})_{\mathbb{L}_2(SE(2), M, d\tilde{g})}}, \tag{34}$$

with $M^{roi}(\mathbf{x}, \theta) = m^{roi}(\mathbf{x})$.

4 Additional Details on the Detection Problems

In this section we describe additional details about the implementation and results of the three detection problems discussed in the main article.

Fig. 3. A selection of positive and negative image patches f_i used in the training of templates.

TABLE 1
Average processing times. For optic nerve head detection (ONH) the average is taken over 1529 images of the TC, MESSIDOR, DRIVE and STARE database. For fovea detection the average is taken over 1408 images of the TC and MESSIDOR database. For pupil detection the average is taken over 1521 images of the BioID database.

	ONH		Fovea		Pupil (left & right)	
	$\mathbb{R}^2$	$SE(2)$	$\mathbb{R}^2$	$SE(2)$	$\mathbb{R}^2$	$SE(2)$
	Timings (ms)					
1. Rescaling	106	106	111	111	0	0
2. $\mathbb{R}^2$-Processing	66	66	64	64	71	71
3. OS Transform	0	108	0	108	0	121
4. $SE(2)$-Processing	0	5	0	5	0	6
5. Template Matching	20	195	19	190	26	116
Total	192	479	195	477	97	313
	Combined Total Timings (ms) - $\mathbb{R}^2$ and $SE(2)$					
	497		501		420	
	Combined Total Timings (ms) - Fovea and ONH					
	730					

TABLE 2
Success rates for optic nerve head detection ($\pm$ standard deviation, number of fails in parenthesis) with varying accuracy requirements in 5-fold cross validation. Maximum distance to ground truth location is expressed in optic disk radius R.

Database (# of images)		Maximum distance to ground truth R/8	R/4	R/2	R	2R
ES (SLO)	(208)	98.05% $\pm$ 2.04% (4)	100.0% $\pm$ 0.00% (0)	100.0% $\pm$ 0.00% (0)	100.0% $\pm$ 0.00% (0)	100.0% $\pm$ 0.00% (0)
TC	(208)	84.19% $\pm$ 4.34% (33)	94.54% $\pm$ 3.51% (11)	99.52% $\pm$ 1.06% (1)	100.0% $\pm$ 0.00% (0)	100.0% $\pm$ 0.00% (0)
MESSIDOR	(1200)	73.07% $\pm$ 3.69% (323)	94.41% $\pm$ 1.47% (67)	99.50% $\pm$ 0.46% (6)	99.92% $\pm$ 0.19% (1)	100.0% $\pm$ 0.00% (0)
DRIVE	(40)	70.84% $\pm$ 26.0% (13)	91.69% $\pm$ 12.3% (4)	98.18% $\pm$ 4.07% (1)	98.18% $\pm$ 4.07% (1)	100.0% $\pm$ 0.00% (0)
STARE	(81)	48.12% $\pm$ 10.27% (42)	74.94% $\pm$ 6.52% (20)	89.39% $\pm$ 8.16% (9)	98.67% $\pm$ 2.98% (1)	98.67% $\pm$ 2.98% (1)
All Images	(1737)	76.11% $\pm$ 2.58% (415)	94.13% $\pm$ 0.79% (102)	99.02% $\pm$ 0.26% (17)	99.83% $\pm$ 0.26% (3)	99.94% $\pm$ 0.13% (1)

4.1 Training Samples

In all three applications training samples were used to compute the templates. Positive training samples were centered around the object of interest. Negative training samples were centered around random locations in the image, but not within a certain distances to the true positive object location. In the retinal applications this distance was one optic disk radius, in the pupil detection application this was a normalized distance of 0.1 (cf. Eq.(39) of the main article). An selection of the 2D image pathes that were used in the experiments are shown in Fig. 3.

4.2 Processing Pipeline, Settings and Timings

4.2.1 Processing Pipeline

In all three application the same processing pipeline was used. The pipeline can be divided into the following 5 steps:

1) *Resizing*. Each input image is resized to a certain operating resolution and cropped to remove large regions with value 0 (outside the field of view mask in retinal images, see e.g. Fig. 4). The retinal images are resized such that the pixel size was approximately $40\mu m/pix$. In the pupil detection application no rescaling or cropping was done.
2) $\mathbb{R}^2$*-Processing*. In all three applications we applied a local intensity and contrast normalization step using an adaptation of [10] which we explain below. The locally normalized image $\hat{f}$ is then mapped through an error function via $\mathrm{erf}(8\hat{f})$ to dampen outliers.
3) *Orientation score transform*. The processed image is then taken as input for an orientation score transform using Eq. (23) of the main article. For the oriented wavelets we used cake wavelets [9], [11] of size $[51\times51]$ and with angular resolution $s_\theta = \pi/12$, and with sampling θ from 0 to π.
4) $SE(2)$*-Processing*. For phase-invariant, nonlinear, left-invariant [5], and contractive [12] processing on SE(2), we work with the modulus of the complex valued orientation scores rather than with the complex-valued scores themselves (taking the modulus of quadrature filter responses is an effective technique for line detection, see e.g. Freeman et al. [13]).
5) *Template Matching*. Finally we perform template matching using respectively Eqs. (3),(4) and (5) of the main article for the $\mathbb{R}^2$ case and Eqs. (3),(25) and (26) of the main article for the $SE(2)$ case.

Regarding the image resolutions (step 1) we note that the average image size after rescaling was $[300 \times 300]$. The average image resolutions in each database were as follows:

- *ES (SLO)* contained images of average resolution $13.9\mu m/pix$.
- *TC* contained images of average resolution $9.4\mu m/pix$.
- *MESSIDOR* contained images of 3 cameras with average resolutions $13.6\mu m/pix$, $9.1\mu m/pix$ and $8.6\mu m/pix$.
- *DRIVE* contained images of average resolution $21.9\mu m/pix$.
- *STARE* contained images of average resolution $17.6\mu m/pix$.

Regarding local image normalization (step 2) we note the following. Local image normalization was done using an adaptation of [10]. The method first computes a local average and standard deviation of pixel intensities, and the image is locally normalized to zero mean and unit standard deviation. This is done via Eq. (31). Then a background mask is construct by setting pixels with a larger distance than 1 standard deviation to the average (Mahalanobis distance) to 0, and other pixels to 1. This mask is then used to ignore outliers in a second computation of the local average and standard deviation. The final normalized image is again computed via Eq. (31) but now with the inclusion of the background mask, see Eq. (33).

4.2.2 Template Settings

In the retinal applications we used $\mathbb{R}^2$ templates of size $[N_x \times N_y] = [251 \times 251]$ which were covered by a grid of B-spline basis functions of size $[N_k \times N_l] = [51\times51]$, the $SE(2)$ templates were of size $[N_x \times N_y \times N_\theta] = [251 \times 251 \times 12]$ and were covered by a grid of B-spline basis functions of size $[N_k \times N_l \times N_m] = [51 \times 51 \times 12]$.

In the pupil detection application we used $\mathbb{R}^2$ templates of size $[N_x \times N_y] = [101 \times 101]$ which were also covered by a grid of B-spline basis functions of size $[N_k \times N_l] = [51 \times 51]$, the $SE(2)$ templates were of size $[N_x \times N_y \times N_\theta] = [101 \times 101 \times 12]$ and were also covered by a grid of B-spline basis functions of size $[N_k \times N_l \times N_m] = [51 \times 51 \times 12]$.

The regularization parameters (λ, μ and $D_{\theta\theta}$) for the different template types were automatically optimized using generalized cross validation.

4.2.3 Timings

We computed the average time for detecting one (or two) object(s) in an image and tabulated the results in Tab. 1. Here we sub-divided the timings into the 5 processing steps explained in Subsec. 4.2.1. The average (full) processing time on the retinal images was in both applications approximately $500ms$. When both the ONH and fovea are detected by the same processing pipeline the processing took $730ms$. For pupil detection the average time to detect *both* the left and right pupil on the *full* images was $420ms$.

The retinal images were on average of size $[1230 \times 1792]$, and $[300 \times 300]$ after cropping and resizing. The images in the pupil detection application were not resized or cropped and were of size $[286 \times 384]$.

All experiments were performed using Wolfram *Mathematica* 10.4, on a computer with an Intel Core i703612QM CPU and 8GB memory.

4.3 Detection Results

In this section we provide the results for the three separate applications. A general discussion of these results can be found in the main article.

4.3.1 Optic Nerve Head Detection

A Table of detection performance for each type of template is provided in Tab. 1 of the main article. In Fig. 4 we show the 3 failed cases for ONH detection, and a selection of correct ONH localizations in difficult images. In Table 2 we show detection results for varying accuracy criteria. Note that detection results are typically reported for the accuracy requirement of 1 optic disk radius with the target (see also state-of-the-art comparison in Table 2 of the main article).

4.3.2 Fovea Detection

A Table of detection performance for each type of template is provided in Tab. 3. In Fig. 5 we show next to a selection of successful detections the only 5 failed cases on images from conventional fundus (CF) cameras (TC, MESSIDOR, DRIVE, STARE), and 3 of the failed detections in images coming from an scanning laser ophthalmoscopy (SLO) camera.

As can also be read from Tab. 3, we found that fovea detection in SLO images was significantly more difficult than fovea detection in CF images. The reason for this is that on SLO images the clear dark blob-like shape is not always present on these images. Compare for example the positive fovea patches from Fig. 3 (where one generally sees a dark blob at the center) with the fovea locations in the bottom row of images in Figs. 4 and 5.

Additionally, the ES (SLO) and CF databases are also more difficult than the MESSIDOR database for fovea detection, as these two databases contain a mix of both fovea centered and ONH centered images. The MESSIDOR database contains only fovea centered images, in which case the fovea is always located around the center of the image. Therefore, even though MESSIDOR is one of the most used databases, it might not be the most representative database for fovea detection benchmarking.

We show detection performance for a range of accuracy requirements in Table 4 for the different databases used in our experiments, and in Table 5 a comparison to the state of the art. There we see that for the stricter requirement of detection within half an optic disk radius our method still outperforms the state of the art. We also see that with further decreasing the acceptance distance ($R/4$ or lower) none of the methods provided acceptable results.

4.3.3 Pupil Detection

A Table of detection performance for each type of template is provided in Tab. 6. In Fig. 6 we show a selection of failed and successful detections. By inspection of the failed cases we found that a main source of failed detections was due to rotations of the head. As stated in the previous section 4.2 we did not employ a rotation invariant detection scheme. Doing so might improve the results. Other failed detections could be attributed to closed eyes, reflection of glasses, distracting background objects and different scales (object distance to camera).

5 Rotation-Scale Invariant Matching

5.1 A Basic Extension

The extension to rotation and scale invariant object localization of the 2D cross-correlation based template matching approach, described in Eqs. (3)-(5) of the main article, is as follows. For the linear potential function (Eq. (4) of the main article) we can define

$$P^{\mathbb{R}^2}_{lin,inv}(\mathbf{x}) := \max_{\substack{a \in [a_-, a_+],\\ \alpha \in [0, 2\pi)}} (\mathcal{T}_{\mathbf{x}} \mathcal{S}_a \mathcal{R}_\alpha \, t, f)_{\mathbb{L}_2(\mathbb{R}^2)}, \quad (35)$$

and for the logistic regression case (Eq. (5) of the main article) we define

$$P^{\mathbb{R}^2}_{log,inv}(\mathbf{x}) := \max_{\substack{a \in [a_-, a_+],\\ \alpha \in [0, 2\pi)}} \sigma\left((\mathcal{T}_{\mathbf{x}} \mathcal{S}_a \mathcal{R}_\alpha \, t, f)_{\mathbb{L}_2(\mathbb{R}^2)}\right), \quad (36)$$

with σ the logistic sigmoid function defined in Eq. (5) of the main article, and with rotation operator $\mathcal{R}_\alpha$ and scaling operator $\mathcal{S}_a$ defined by

$$(\mathcal{R}_\alpha t)(\mathbf{x}) = t(\mathbf{R}_\alpha^{-1}\mathbf{x}), \quad (37)$$

$$(\mathcal{S}_a t)(\mathbf{x}) = a^{-1} t(a\mathbf{x}), \quad (38)$$

with rotation matrix $\mathbf{R}_\alpha$ representing a counter clockwise rotation of angle α. By taking the maximum over scales a (in a suitable range $[a_-, a_+]$) and rotations α, the response of the best matching template is obtained at each location $\mathbf{x}$, and invariance is obtained with respect to scaling and rotation of the object of interest.

The rotation/scale invariant extension of the $SE(2)$ cross-correlation template matching case (Eqs. (25)-(26) of the main article) is equally straightforward; for the linear potential we define

$$P^{SE(2)}_{lin,inv}(\mathbf{x}) := \max_{\substack{a \in [a_-, a_+],\\ \alpha \in [0, 2\pi)}} (\mathcal{T}_{\mathbf{x}} \mathcal{S}_a \mathcal{R}_\alpha \, T, U_f)_{\mathbb{L}_2(SE(2))}, \quad (39)$$

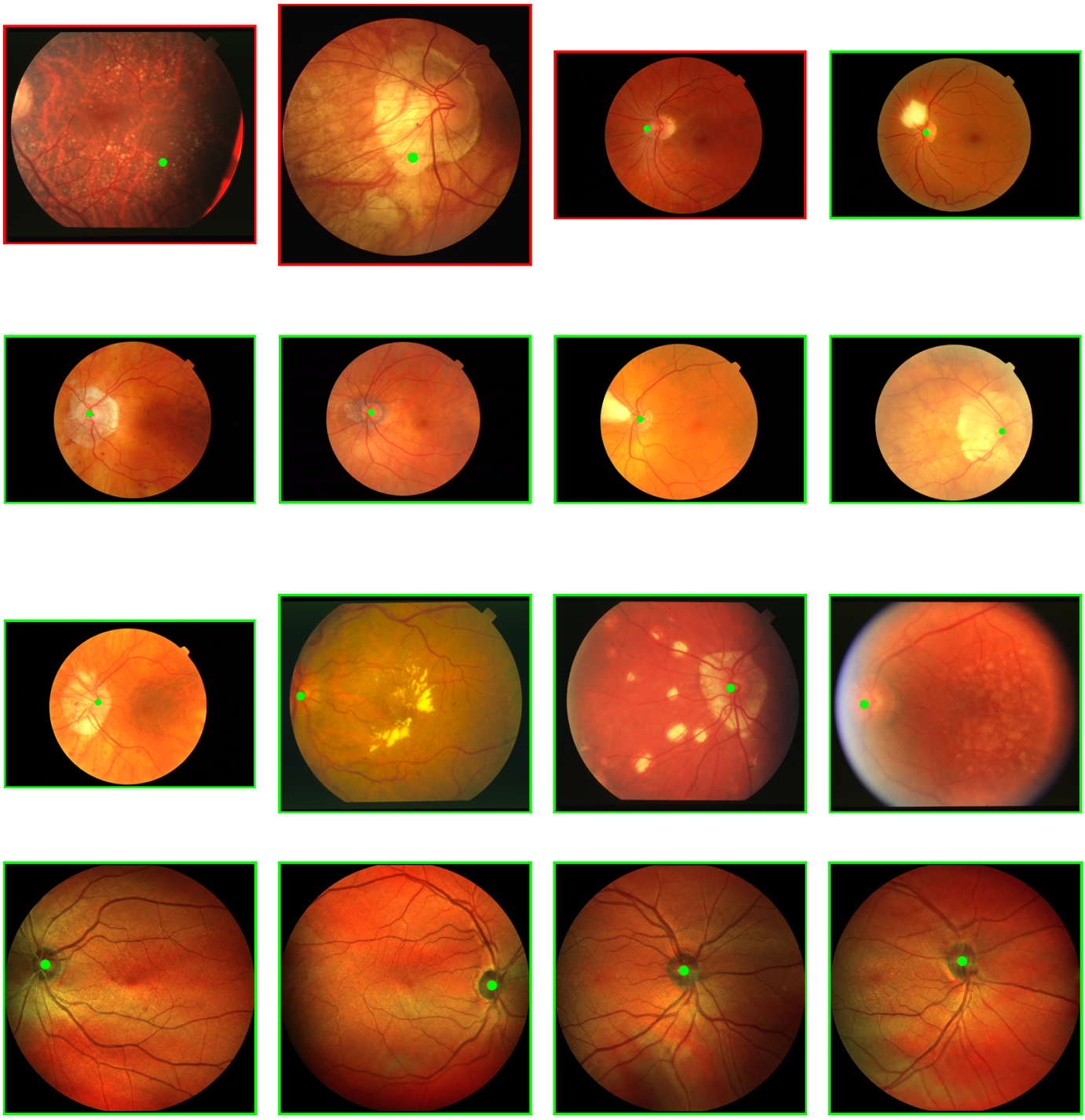

Fig. 4. Detection results of our best method for optic nerve head detection in retinal images. Successful detection are indicated with a green frame around the image, failed detections are indicated with a red frame. In the ONH detection application there were only 3 fails in a set of 1737 images.

TABLE 3
Average template matching results ($\pm$ standard deviation) for fovea detection in 5-fold cross validation, number of failed detections in parentheses.

Template ID	ES (SLO) 208	TC 208	MESSIDOR 1200	All Images 1616
		$\mathbb{R}^2$ templates		
$A_{\mathbb{R}^2}$	76.36% $\pm$ 6.79% (49)	98.24% $\pm$ 2.74% (3)	98.41% $\pm$ 0.22% (19)	95.60% $\pm$ 0.98% (71)
$B_{lin:\mathbb{R}^2}$	23.50% $\pm$ 3.81% (159)	31.66% $\pm$ 9.03% (142)	51.19% $\pm$ 5.97% (587)	45.07% $\pm$ 3.33% (888)
$C_{lin:\mathbb{R}^2}$	45.65% $\pm$ 8.61% (113)	98.24% $\pm$ 2.74% (3)	98.59% $\pm$ 0.36% (17)	91.77% $\pm$ 1.26% (133)
$D_{lin:\mathbb{R}^2}$	44.21% $\pm$ 4.62% (116)	99.49% $\pm$ 1.14% (1)	98.84% $\pm$ 0.31% (14)	91.90% $\pm$ 0.59% (131)
$E_{lin:\mathbb{R}^2}$	46.10% $\pm$ 8.11% (112)	98.86% $\pm$ 1.57% (2)	98.67% $\pm$ 0.34% (16)	91.95% $\pm$ 1.18% (130)
$B_{log:\mathbb{R}^2}$	1.43% $\pm$ 1.31% (205)	10.27% $\pm$ 5.09% (185)	20.07% $\pm$ 3.00% (959)	16.53% $\pm$ 2.52% (1349)
$C_{log:\mathbb{R}^2}$	9.59% $\pm$ 3.74% (188)	70.30% $\pm$ 8.57% (61)	77.61% $\pm$ 4.64% (267)	68.06% $\pm$ 3.53% (516)
$D_{log:\mathbb{R}^2}$	11.48% $\pm$ 4.70% (184)	83.47% $\pm$ 7.80% (32)	88.22% $\pm$ 2.81% (141)	77.90% $\pm$ 2.00% (357)
$E_{log:\mathbb{R}^2}$	2.86% $\pm$ 2.62% (202)	79.68% $\pm$ 7.92% (40)	84.79% $\pm$ 5.16% (181)	73.82% $\pm$ 2.62% (423)
		$SE(2)$ templates		
$A_{SE(2)}$	67.81% $\pm$ 4.69% (67)	79.13% $\pm$ 9.11% (40)	98.25% $\pm$ 0.68% (21)	92.08% $\pm$ 0.84% (128)
$B_{lin:SE(2)}$	83.19% $\pm$ 2.76% (35)	71.53% $\pm$ 7.36% (58)	91.31% $\pm$ 0.68% (104)	87.81% $\pm$ 1.25% (197)
$C_{lin:SE(2)}$	83.65% $\pm$ 3.18% (34)	84.13% $\pm$ 6.25% (32)	98.23% $\pm$ 1.04% (21)	94.62% $\pm$ 0.36% (87)
$D_{lin:SE(2)}$	73.57% $\pm$ 4.71% (55)	83.69% $\pm$ 6.83% (33)	97.88% $\pm$ 1.17% (25)	93.01% $\pm$ 1.09% (113)
$E_{lin:SE(2)}$	77.83% $\pm$ 4.29% (46)	84.88% $\pm$ 6.69% (30)	98.22% $\pm$ 1.23% (21)	94.00% $\pm$ 0.93% (97)
$B_{log:SE(2)}$	75.49% $\pm$ 5.73% (51)	60.80% $\pm$ 5.68% (80)	92.79% $\pm$ 1.98% (86)	86.56% $\pm$ 2.20% (217)
$C_{log:SE(2)}$	79.33% $\pm$ 6.57% (43)	70.87% $\pm$ 10.28% (59)	96.90% $\pm$ 0.71% (37)	91.39% $\pm$ 1.36% (139)
$D_{log:SE(2)}$	62.09% $\pm$ 6.66% (79)	72.57% $\pm$ 8.59% (54)	96.64% $\pm$ 1.05% (40)	89.30% $\pm$ 0.63% (173)
$E_{log:SE(2)}$	68.34% $\pm$ 8.59% (66)	72.20% $\pm$ 8.53% (55)	96.57% $\pm$ 0.96% (41)	89.98% $\pm$ 1.25% (162)
		Template combinations (sorted on performance)		
$C_{lin:\mathbb{R}^2} + C_{log:SE(2)}$	97.17% $\pm$ 3.01% (6)	99.17% $\pm$ 1.13% (2)	99.74% $\pm$ 0.38% (3)	99.32% $\pm$ 0.26% (11)
* $A_{\mathbb{R}^2} + C_{lin:SE(2)}$	98.08% $\pm$ 2.03% (4)	98.07% $\pm$ 1.95% (4)	99.68% $\pm$ 0.33% (4)	99.26% $\pm$ 0.47% (12)
$E_{lin:\mathbb{R}^2} + C_{log:SE(2)}$	96.20% $\pm$ 3.15% (8)	99.17% $\pm$ 1.13% (2)	99.75% $\pm$ 0.23% (3)	99.20% $\pm$ 0.35% (13)
$E_{lin:\mathbb{R}^2} + C_{lin:SE(2)}$	96.65% $\pm$ 2.13% (7)	99.17% $\pm$ 1.13% (2)	99.66% $\pm$ 0.36% (4)	99.19% $\pm$ 0.42% (13)
$C_{lin:\mathbb{R}^2} + C_{lin:SE(2)}$	97.14% $\pm$ 1.97% (6)	98.78% $\pm$ 1.78% (3)	99.58% $\pm$ 0.31% (5)	99.13% $\pm$ 0.40% (14)
$A_{\mathbb{R}^2} + E_{lin:SE(2)}$	97.59% $\pm$ 1.73% (5)	98.07% $\pm$ 1.95% (4)	99.59% $\pm$ 0.28% (5)	99.13% $\pm$ 0.25% (14)
$E_{lin:\mathbb{R}^2} + E_{lin:SE(2)}$	96.16% $\pm$ 2.76% (8)	99.17% $\pm$ 1.13% (2)	99.58% $\pm$ 0.31% (5)	99.07% $\pm$ 0.38% (15)
$E_{lin:\mathbb{R}^2} + D_{lin:SE(2)}$	95.71% $\pm$ 3.07% (9)	99.17% $\pm$ 1.13% (2)	99.58% $\pm$ 0.31% (5)	99.01% $\pm$ 0.40% (16)
$C_{lin:\mathbb{R}^2} + E_{lin:SE(2)}$	96.16% $\pm$ 2.76% (8)	98.78% $\pm$ 1.78% (3)	99.58% $\pm$ 0.31% (5)	99.01% $\pm$ 0.51% (16)
$A_{\mathbb{R}^2} + C_{log:SE(2)}$	96.65% $\pm$ 2.13% (7)	98.07% $\pm$ 1.95% (4)	99.58% $\pm$ 0.42% (5)	99.01% $\pm$ 0.26% (16)
...		...		...
† $A_{\mathbb{R}^2} + A_{SE(2)}$	92.85 % $\pm$ 4.68% (15)	95.84% $\pm$ 2.58% (8)	99.58% $\pm$ 0.30% (5)	98.27% $\pm$ 0.70% (28)
...		...		...

Best template combination that does not rely on logistic regression.
†*Best template combination that does not rely on template optimization.*

TABLE 4
Success rates for fovea detection ($\pm$ standard deviation, number of fails in parenthesis) with varying accuracy requirements in 5-fold cross validation. Maximum distance to ground truth location is expressed in optic disk radius R.

Database (# of images)		Maximum distance to ground truth R/8	R/4	R/2	R	2R
ES (SLO)	(208)	66.91% $\pm$ 4.64% (69)	92.85% $\pm$ 3.16% (15)	94.74% $\pm$ 1.93% (11)	97.17% $\pm$ 3.01% (6)	97.66% $\pm$ 3.28% (5)
TC	(208)	49.51% $\pm$ 4.07% (106)	80.33% $\pm$ 3.22% (40)	95.41% $\pm$ 1.77% (9)	99.17% $\pm$ 1.13% (2)	99.61% $\pm$ 0.88% (1)
MESSIDOR	(1200)	61.81% $\pm$ 2.64% (459)	90.56% $\pm$ 1.31% (113)	98.07% $\pm$ 0.87% (23)	99.74% $\pm$ 0.38% (3)	100.0% $\pm$ 0.00% (0)
All Images	(1616)	60.78% $\pm$ 1.84% (634)	89.60% $\pm$ 0.80% (168)	97.34% $\pm$ 0.65% (43)	99.32% $\pm$ 0.26% (11)	99.63% $\pm$ 0.40% (6)

TABLE 5
Success rates for fovea detection (number of fails in parenthesis) with varying accuracy requirements; a comparison to literature using the MESSIDOR database. Maximum distance to ground truth location is expressed in optic disk radius R.

Method	Maximum distance to ground truth				
	R/8	R/4	R/2	R	2R
Niemeijer et al. [14], [15]	75.67% (292)	93.50% (78)	96.83% (38)	97.92% (25)	-
Yu et al. et al. [16]	-	-	95.00% (60)	-	-
Gegundez-Arias et al. [15]	80.42% (235)	93.90% (73)	96.08% (47)	96.92% (37)	97.83% (26)
Giachetti et al. [17]	-	-	-	99.10% (11)	-
Aquino [18]	-	-	-	98.20% (21)	-
Proposed	61.81% (459)	90.56% (113)	98.07% (23)	99.74% (3)	100.0% (0)

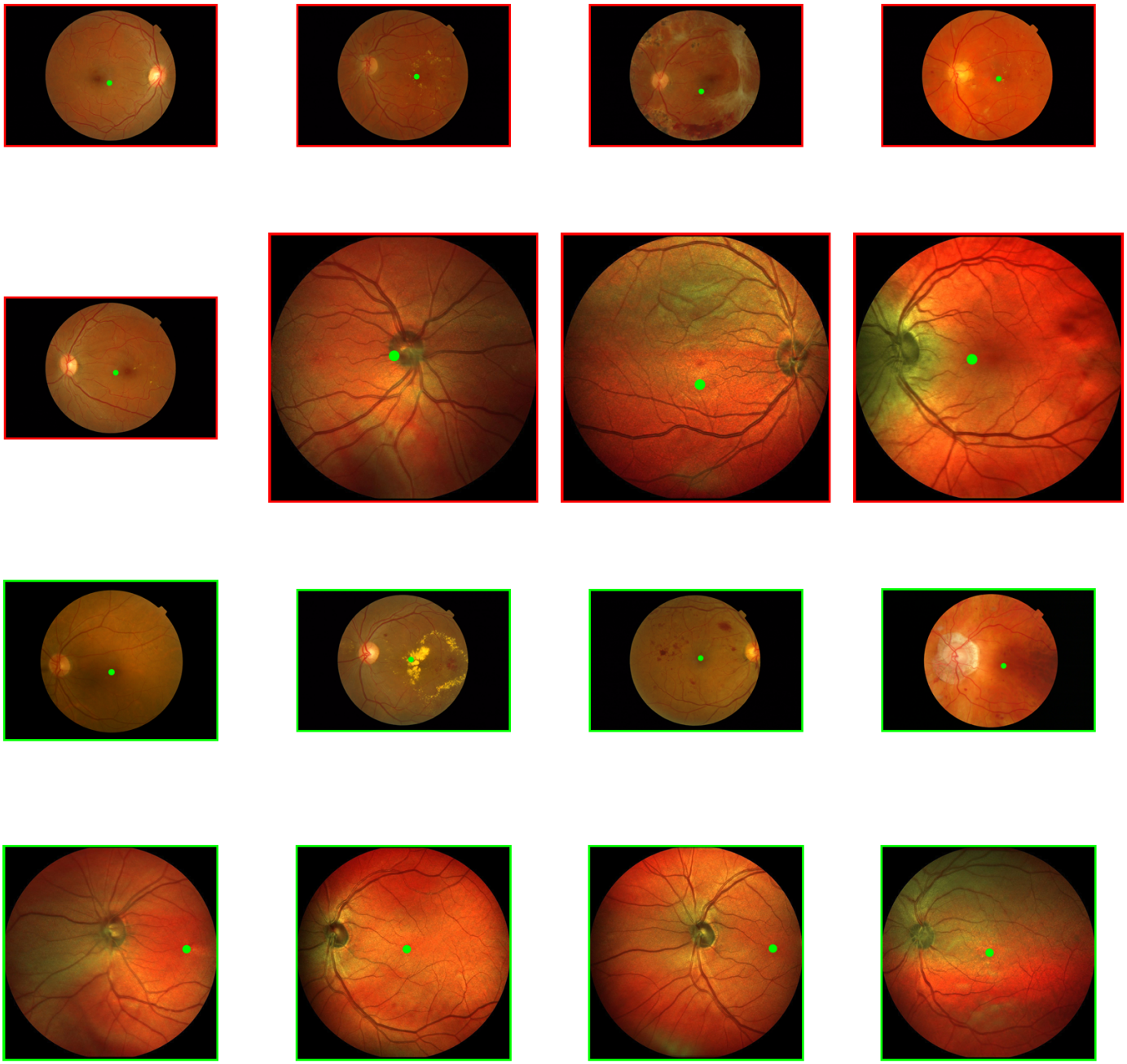

Fig. 5. Detection results of our best method for fovea detection in retinal images. Successful detection are indicated with a green frame around the image, failed detections are indicated with a red frame. In the fovea detection application there were only 5 fails in a set of 1408 conventional fundus (CF) camera images. Out of the 208 scanning laser ophthalmoscopy (SLO) images there were 6 fails, 3 of them are shown in this figure.

Fig. 6. Detection results of our best method for pupil detection. Successful detection are indicated with a green frame around the image, failed detections are indicated with a red frame.

and for the logistic potential we define

$$P^{SE(2)}_{log,inv}(\mathbf{x}) := \max_{\substack{a \in [a_-, a_+], \\ \alpha \in [0, 2\pi)}} \sigma\left((\mathcal{T}_{\mathbf{x}}\mathcal{S}_a\mathcal{R}_\alpha\, T, U_f)_{\mathbb{L}_2(SE(2))}\right), \tag{40}$$

with for orientation score objects $T, U_f \in \mathbb{L}_2(SE(2))$ the rotation and scaling operators defined respectively by

$$(\mathcal{R}_\alpha T)(\mathbf{x}, \theta) = T(\mathbf{R}_\alpha^{-1}\mathbf{x}, \theta - \alpha), \tag{41}$$

$$(\mathcal{S}_a T)(\mathbf{x}, \theta) = a^{-1} T(a\mathbf{x}, \theta). \tag{42}$$

It depends on the addressed template matching problem whether or not such invariance is desirable or not. In many applications the object is to be found in a human environment context, in which some objects tend to appear in specific orientations or at typical scales, and in which case rotation/scale invariance might not be desirable. E.g. the sizes of anatomical structures in the retina are relatively constant among different subjects (constant scale) and retinal images are typically taken at a fixed orientation (constant rotation). In the pupil detection problem the subjects typically appear in upright position behind the camera (constant rotation), and within a reasonable distance to the camera (constant scale). In the next Subsec. 5.2.1 we indeed show that in the applications considered in this manuscript rotation/scale invariance is not necessarily a desired property, and that computation time linearly increases with the number of rotations/scalings tested for (cf. Subsec. 5.2.2).

5.2 Results with Rotation and Scale Invariance

Here we perform rotation and scale invariant template matching via the extension described in Subsec. 5.1. We selected the best template combination for each specific application and compared non-invariant template matching (as described in the main article) to rotation and/or scale invariant template matching (Subsec. 5.1). The best template combination for ONH detection was $A_{\mathbb{R}^2} + C_{log:SE(2)}$, for fovea detection this was $C_{lin:\mathbb{R}^2} + C_{log:SE(2)}$, and for pupil detection this was $D_{lin:\mathbb{R}^2} + E_{lin:SE(2)}$.

For the retinal applications we only tested for rotation invariance with

$$\alpha \in \{-\frac{\pi}{6}, -\frac{\pi}{8}, -\frac{\pi}{12}, -\frac{\pi}{24}, 0, \frac{\pi}{24}, \frac{\pi}{12}, \frac{\pi}{8}, \frac{\pi}{6}\},$$

and did not included scale invariance since each retinal image was already rescaled to a standardized resolution (see

TABLE 6
Average template matching results ($\pm$ standard deviation) for pupil detection in 5-fold cross validation, number of failed detections in parentheses. A successful detection has a normalized error $e \leq 0.1$.

Template ID	BioID (Full image) 1521	BioID (Periocular image) 1521
$\mathbb{R}^2$ templates		
$A_{\mathbb{R}^2}$	41.03% ± 1.45% (897)	59.70% ± 1.52% (613)
$B_{lin:\mathbb{R}^2}$	0.00% ± 0.00% (1521)	3.62% ± 1.09% (1466)
$C_{lin:\mathbb{R}^2}$	12.95% ± 2.22% (1324)	67.26% ± 2.55% (498)
$D_{lin:\mathbb{R}^2}$	8.28% ± 1.80% (1395)	75.68% ± 2.33% (370)
$E_{lin:\mathbb{R}^2}$	11.51% ± 2.25% (1346)	71.47% ± 2.76% (434)
$B_{log:\mathbb{R}^2}$	0.00% ± 0.00% (1521)	0.00% ± 0.00% (1521)
$C_{log:\mathbb{R}^2}$	12.89% ± 2.06% (1325)	39.91% ± 3.37% (914)
$D_{log:\mathbb{R}^2}$	1.84% ± 0.95% (1493)	22.09% ± 2.37% (1185)
$E_{log:\mathbb{R}^2}$	10.39% ± 2.26% (1363)	37.21% ± 4.37% (955)
$SE(2)$ templates		
$A_{SE(2)}$	57.72% ± 1.68% (643)	75.34% ± 1.31% (375)
$B_{lin:SE(2)}$	8.74% ± 2.00% (1388)	41.81% ± 5.04% (885)
$C_{lin:SE(2)}$	84.61% ± 4.19% (234)	86.78% ± 3.68% (201)
$D_{lin:SE(2)}$	85.53% ± 3.44% (220)	87.18% ± 3.71% (195)
$E_{lin:SE(2)}$	85.47% ± 3.82% (221)	87.11% ± 3.87% (196)
$B_{log:SE(2)}$	0.00% ± 0.00% (1521)	0.13% ± 0.29% (1519)
$C_{log:SE(2)}$	86.52% ± 0.77% (205)	93.95% ± 1.33% (92)
$D_{log:SE(2)}$	75.21% ± 2.18% (377)	89.48% ± 2.27% (160)
$E_{log:SE(2)}$	83.30% ± 1.68% (254)	92.77% ± 1.02% (110)
Template combinations (sorted on performance full image)		
$^{*}C_{lin:\mathbb{R}^2} + E_{lin:SE(2)}$	93.49% ± 1.49% (99)	95.60% ± 1.46% (67)
$C_{lin:\mathbb{R}^2} + D_{lin:SE(2)}$	93.16% ± 1.54% (104)	95.00% ± 1.15% (76)
$E_{lin:\mathbb{R}^2} + E_{lin:SE(2)}$	93.10% ± 1.04% (105)	95.59% ± 0.89% (67)
$E_{lin:\mathbb{R}^2} + D_{lin:SE(2)}$	92.97% ± 1.62% (107)	95.27% ± 1.31% (72)
$C_{lin:\mathbb{R}^2} + C_{lin:SE(2)}$	92.70% ± 1.41% (111)	95.33% ± 0.97% (71)
$E_{lin:\mathbb{R}^2} + C_{lin:SE(2)}$	92.64% ± 0.94% (112)	95.33% ± 0.94% (71)
$D_{lin:\mathbb{R}^2} + D_{lin:SE(2)}$	92.51% ± 0.96% (114)	95.79% ± 0.82% (64)
$D_{lin:\mathbb{R}^2} + E_{lin:SE(2)}$	92.24% ± 1.23% (118)	95.86% ± 0.89% (63)
$E_{log:\mathbb{R}^2} + D_{lin:SE(2)}$	92.11% ± 2.26% (120)	93.23% ± 1.93% (103)
$D_{lin:\mathbb{R}^2} + C_{log:SE(2)}$	92.05% ± 1.52% (121)	95.14% ± 0.78% (74)
...	...	
Template combinations (sorted on performance periocular image)		
$^{*}D_{lin:\mathbb{R}^2} + E_{lin:SE(2)}$	92.24% ± 1.23% (118)	95.86% ± 0.89% (63)
$D_{lin:\mathbb{R}^2} + D_{lin:SE(2)}$	92.51% ± 0.96% (114)	95.79% ± 0.82% (64)
$D_{lin:\mathbb{R}^2} + C_{lin:SE(2)}$	91.52% ± 1.25% (129)	95.73% ± 0.77% (65)
$E_{lin:\mathbb{R}^2} + E_{lin:SE(2)}$	93.10% ± 1.04% (105)	95.59% ± 0.89% (67)
$C_{lin:\mathbb{R}^2} + E_{lin:SE(2)}$	93.49% ± 1.49% (99)	95.60% ± 1.46% (67)
$E_{lin:\mathbb{R}^2} + C_{lin:SE(2)}$	92.64% ± 0.94% (112)	95.33% ± 0.94% (71)
$C_{lin:\mathbb{R}^2} + C_{lin:SE(2)}$	92.70% ± 1.41% (111)	95.33% ± 0.97% (71)
$E_{lin:\mathbb{R}^2} + D_{lin:SE(2)}$	92.97% ± 1.62% (107)	95.27% ± 1.31% (72)
$D_{lin:\mathbb{R}^2} + E_{log:SE(2)}$	91.72% ± 1.23% (126)	95.27% ± 0.79% (72)
$D_{lin:\mathbb{R}^2} + C_{log:SE(2)}$	92.05% ± 1.52% (121)	95.14% ± 0.78% (74)
...	...	
$^{\dagger}A_{\mathbb{R}^2} + A_{SE(2)}$	61.34% ± 1.54% (588)	68.18% ± 1.25% (484)
...	...	

*Best template combination that does not rely on logistic regression.
†Best template combination that does not rely on template optimization.

TABLE 7
Average template matching results ($\pm$ standard deviation, number of fails between parenthesis) for optic nerve head (ONH), fovea, and pupil detection in 5-fold cross validation.

Method	Success rate
ONH Detection	(1737 images)
No invariance	99.83% ± 0.26% (3)
Rotation invariance	99.60% ± 0.16% (7)
Fovea Detection	(1616 images)
No invariance	99.32% ± 0.26% (11)
Rotation invariance	97.10% ± 0.65% (47)
Pupil Detection	(1521 images)
No invariance	95.86% ± 0.89% (63)
Rotation invariance	94.48% ± 1.62% (84)
Scale invariance	95.33% ± 1.46% (71)
Rotation + scale invariance	94.28% ± 2.10% (87)

Subsec. 4.2.1). In pupil detection we tested for a range of scalings with

$$a \in \{0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3\}$$

to deal with varying pupil sizes caused by varying distances to the camera; and we tested for a range of rotations with

$$\alpha \in \{-\frac{\pi}{4}, -\frac{\pi}{8}, -\frac{\pi}{16}, 0, \frac{\pi}{16}, \frac{\pi}{8}, \frac{\pi}{4}\}$$

to deal with rotations of the head.

5.2.1 Detection Results

The detection results are shown in Table. 7. Here we can see that in all three applications the inclusion of a rotation/scale invariant matching scheme results in a slight decrease in performance. This can be explained by the fact that variations in scale an rotation within the databases are small, and that the trained templates can already deal robustly with these variations (due to the presence of such variations in the training set). By introducing rotation/scale invariance one then only increases the likelihood of false positive detections.

5.2.2 Computation Time

The effect on computation time of rotation/scale invariant matching is shown in Fig. 7. Here one sees that computation time linearly increases with the number of template rotations and scalings tested for. This timings-experiment is performed on the pupil detection application, and the shown timings are only of *step 5* of the full detection pipeline (see Subsec. 4.2.1 and Table 1) as this is the only step that is affected by the rotation/scale invariant extension.

REFERENCES

[1] G. Citti and A. Sarti, "A cortical based model of perceptual completion in the roto-translation space," *JMIV*, vol. 24, no. 3, pp. 307–326, 2006.

[2] J. Zhang, R. Duits, G. Sanguinetti, and B. M. ter Haar Romeny, "Numerical approaches for linear left-invariant diffusions on se(2), their comparison to exact solutions, and their applications in retinal imaging," *Numerical Mathematics: Theory, Methods and Applications*, vol. 9, pp. 1–50, 2 2016.

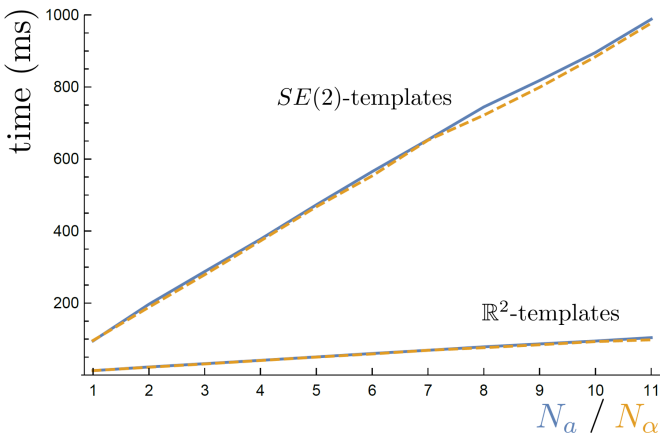

Fig. 7. Average computation times for the detection of one pupil/eye per image, using $SE(2)$ or $\mathbb{R}^2$ templates, testing for different template orientations or scalings. Two experiments are shown, in blue the number of template orientations $N_\alpha = 1$ and the number of scalings N_a is varied, in orange-dashed the number of scalings $N_a = 1$ and the number of rotations is N_α is varied.

[3] R. Duits and B. Burgeth, "Scale spaces on lie groups," in *Scale Space and Variational Methods in Computer Vision*, ser. Lecture Notes in Computer Science, F. Sgallari, A. Murli, and N. Paragios, Eds. Springer Berlin Heidelberg, 2007, vol. 4485, pp. 300–312.

[4] K. Yosida, "Resolvent and spectrum," in *Functional Analysis*, ser. Classics in Mathematics. Springer Berlin Heidelberg, 1995, vol. 123, pp. 209–231.

[5] R. Duits and E. Franken, "Left-invariant parabolic evolutions on SE(2) and contour enhancement via invertible orientation scores part I: Linear left-invariant diffusion equations on SE(2)," *Quarterly of Applied Mathematics*, vol. 68, no. 2, pp. 255–292, 2010.

[6] L. Hörmander, "Hypoelliptic second order differential equations," *Acta Mathematica*, vol. 119, no. 1, pp. 147–171, 1967.

[7] M. Unser, A. Aldroubi, M. Eden, and L. Fellow, "B-spline signal processing: Part i-theory," *IEEE TSP*, vol. 41, pp. 821–833, 1993.

[8] E. P. Hsu, *Stochastic analysis on manifolds*. American Mathematical Soc., 2002, vol. 38.

[9] E. Bekkers, R. Duits, T. Berendschot, and B. ter Haar Romeny, "A multi-orientation analysis approach to retinal vessel tracking," *Journal of Mathematical Imaging and Vision*, vol. 49, no. 3, pp. 583–610, 2014.

[10] M. Foracchia *et al.*, "Luminosity and contrast normalization in retinal images." *MEDIA*, vol. 9, no. 3, pp. 179–90, 2005.

[11] R. Duits, M. Felsberg, G. H. Granlund, and B. M. ter Haar Romeny, "Image analysis and reconstruction using a wavelet transform constructed from a reducible representation of the euclidean motion group," *IJCV*, vol. 72, no. 1, pp. 79–102, 2007.

[12] J. Bruna and S. Mallat, "Invariant scattering convolution networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1872–1886, 2013.

[13] W. Freeman and E. Adelson, "The design and use of steerable filters," *IEEE TPAMI*, vol. 13, no. 9, pp. 891–906, Sep 1991.

[14] M. Niemeijer, M. D. Abràmoff, and B. van Ginneken, "Fast detection of the optic disc and fovea in color fundus photographs." *MEDIA*, vol. 13, no. 6, pp. 859–70, 2009.

[15] M. E. Gegundez-Arias, D. Marin *et al.*, "Locating the fovea center position in digital fundus images using thresholding and feature extraction techniques," *Computerized Medical Imaging and Graphics*, vol. 37, no. 56, pp. 386 – 393, 2013, retinal Image Analysis.

[16] H. Yu, S. Barriga, C. Agurto, S. Echegaray *et al.*, "Fast localization of optic disc and fovea in retinal images for eye disease screening," *Proc. SPIE*, vol. 7963, pp. 796 317–796 317–12, 2011.

[17] A. Giachetti, L. Ballerini, E. Trucco, and P. Wilson, "The use of radial symmetry to localize retinal landmarks," *Computerized Medical Imaging and Graphics*, vol. 37, no. 56, pp. 369 – 376, 2013, retinal Image Analysis.

[18] A. Aquino, "Establishing the macular grading grid by means of fovea centre detection using anatomical-based and visual-based features," *Comp. in Biology and Medicine*, vol. 55, pp. 61 – 73, 2014.