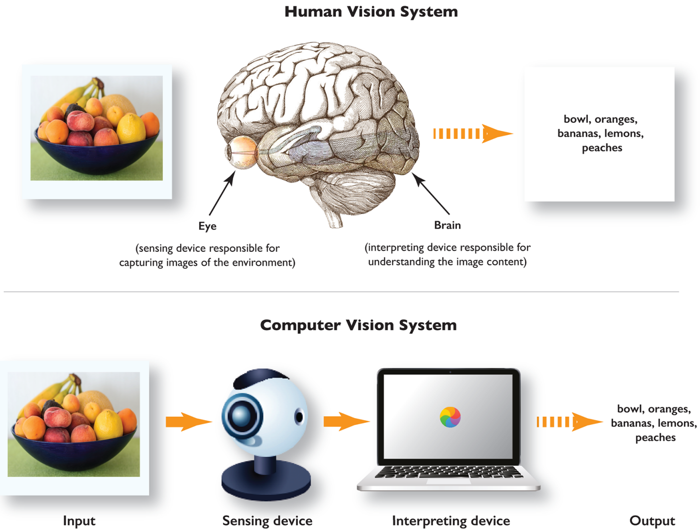
Homografie

przekształcenia w kontekście obrazów

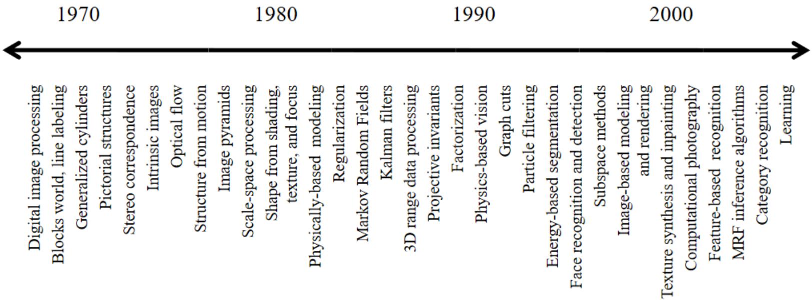
Agata Cieślik

12 maja 2021 r.

Computer vision

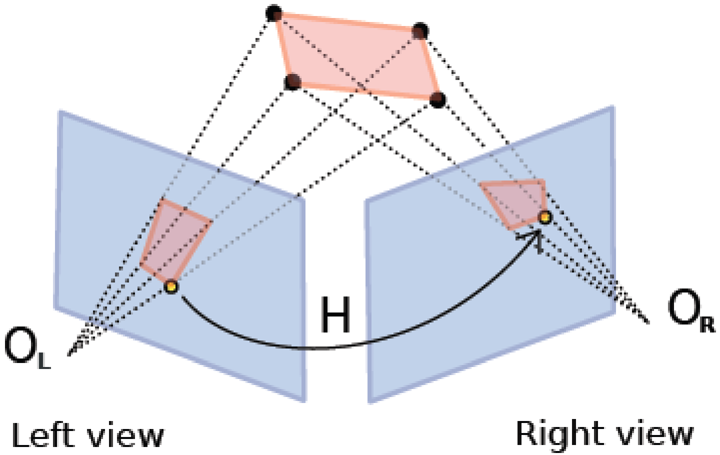


Oś czasu



1970
1980
1990
2000
Digital image processing
Blocks world, line labeling
Generalized cylinders
Pictorial structures
Stereo correspondence
Intrinsic images
Optical flow
Structure from motion
Image pyramids
Scale-space processing
Shape from shading, texture, and focus
Physically-based modeling
Regularization
Markov Random Fields
Kalman filters
3D range data processing
Projective invariants
Factorization
Physics-based vision
Graph cuts
Particle filtering
Energy-based segmentation
Face recognition and detection
Subspace methods
Image-based modeling and rendering
Texture synthesis and inpainting
Computational photography
Feature-based recognition
MRF inference algorithms
Category recognition
Learning

Homografia



Założenia

Obraz możemy traktować jako zbiór pikseli. Piksele natomiast możemy traktować jako dwuwymiarowe punkty $x = (x, y) \in \mathbb{R}^2$ o współrzędnych określających położenie względem krawędzi obrazu.
Przekształcenia obrazów rozważamy więc jako przekształcenia przestrzeni dwuwymiarowej.

Współrzędne jednorodne

W zastosowaniach rozpoznawania obrazów zwykle stosuje się współrzędne jednorodne (*homogenous coordinates*).
Współrzędne jednorodne to sposób reprezentacji punktow n-wymiarowej przestrzeni rzutowej za pomocą układu $n+1$ współrzędnych.
Współrzędne w przestrzeni n wymiarowej uzupełnia się o możliwe kierunki, np. dla punktu dwuwymiarowego x:

- współrzędne kartezjańskie: $x = (x, y) \in \mathbb{R}^2$
- współrzędne jednorodne: $\tilde{x} = (\tilde{x}, \tilde{y}, \tilde{w}) \in \mathcal{P}^2 = \mathbb{R}^3 - (0, 0, 0)$

W szczególności często wykorzystywana jest reprezentacja $\overline{x} = (x, y, 1)$ (*augmented vector*).

Współrzędne jednorodne: dlaczego są użyteczne?

1. Według współrzędnych jednorodnych obrazy różnicące się jedynie skalą są sobie równoważne.
2. Współrzędne jednorodne upraszczają obliczanie wielokrotnych przekształceń obrazu: nie wymagają nakładania translacji przez dodanie osobnego wektora

Współrzędne jednorodne: dlaczego są użyteczne?

Nakładanie przekształceń:

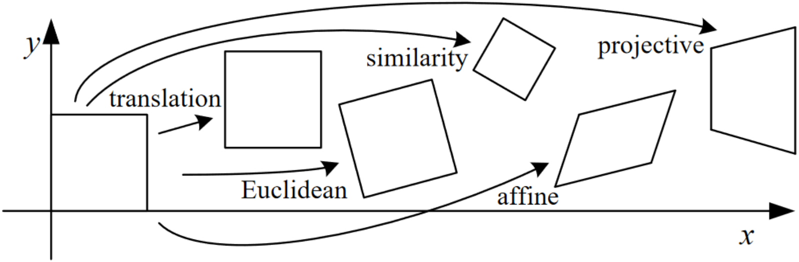
- z użyciem współrzędnych kartezjańskich

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = A \cdot X + T = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} = \begin{bmatrix} a_{11}x + a_{12}y + t_x \\ a_{21}x + a_{22}y + t_y \end{bmatrix}$$

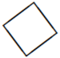
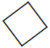
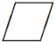
- z użyciem współrzędnych jednorodnych

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = A \cdot X = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} a_{11}x + a_{12}y + t_x w \\ a_{21}x + a_{22}y + t_y w \\ w \end{bmatrix}$$

Przekształcenia



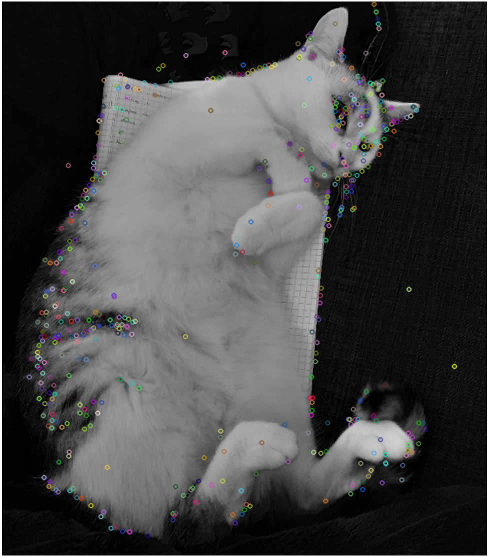
Przekształcenia

Transformation	Matrix	# DoF	Preserves	Icon
translation	$\begin{bmatrix} \boldsymbol{I} \mid \boldsymbol{t} \end{bmatrix}_{2\times 3}$	2	orientation	
rigid (Euclidean)	$\begin{bmatrix} \boldsymbol{R} \mid \boldsymbol{t} \end{bmatrix}_{2\times 3}$	3	lengths	
similarity	$\begin{bmatrix} s\boldsymbol{R} \mid \boldsymbol{t} \end{bmatrix}_{2\times 3}$	4	angles	
affine	$\begin{bmatrix} \boldsymbol{A} \end{bmatrix}_{2\times 3}$	6	parallelism	
projective	$\begin{bmatrix} \tilde{\boldsymbol{H}} \end{bmatrix}_{3\times 3}$	8	straight lines	

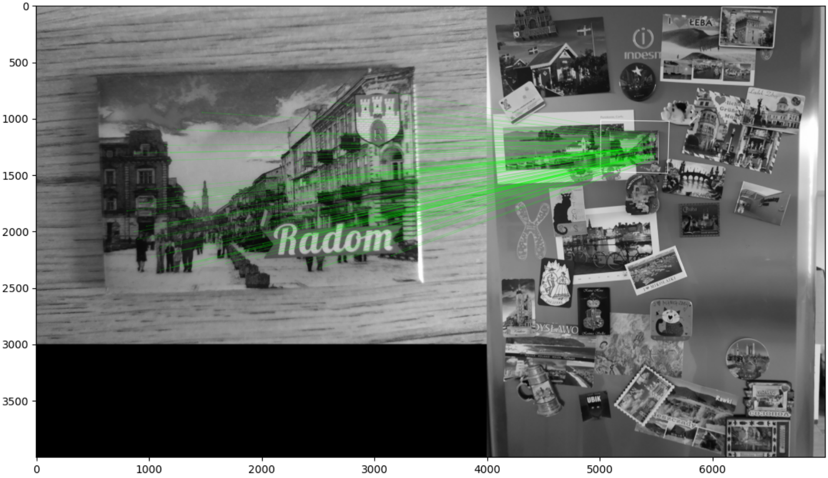
Zastosowanie przekształceń

Jednym z najważniejszych zastosowań jest odnajdywanie cech obiektu na różnych obrazach i określanie relacji między dwoma obrazami.
Przy danych punktach z dwóch obrazów próbuje się wyestymować przekształcenie, które łączy oba obrazy. W przypadku rozpoznawania tego samego obiektu na różnych obrazach zwykle nie będziemy używać do tego celu wszystkich punktów, skupiając się jedynie na tzw. *keypointach* - punktach kluczowych wyznaczonych przez algorytmy typu SIFT, SURF itp.

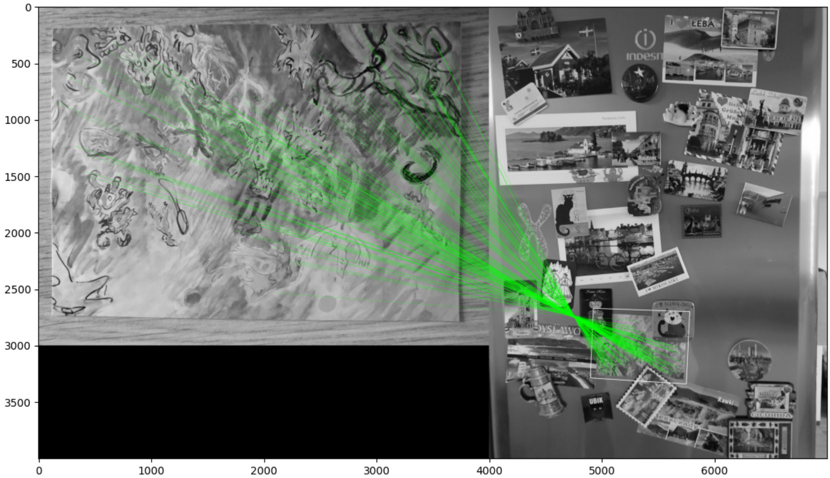
Keypointy



Rozpoznanie obiektu



Rozpoznanie obiektu



Estymacja przekształcenia (homografii)

Transform	Matrix	Parameters p
translation	$\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \end{bmatrix}$	(t_x, t_y)
Euclidean	$\begin{bmatrix} c_\theta & -s_\theta & t_x \\ s_\theta & c_\theta & t_y \end{bmatrix}$	(t_x, t_y, θ)
similarity	$\begin{bmatrix} 1+a & -b & t_x \\ b & 1+a & t_y \end{bmatrix}$	(t_x, t_y, a, b)
affine	$\begin{bmatrix} 1+a_{00} & a_{01} & t_x \\ a_{10} & 1+a_{11} & t_y \end{bmatrix}$	$(t_x, t_y, a_{00}, a_{01}, a_{10}, a_{11})$
projective	$\begin{bmatrix} 1+h_{00} & h_{01} & h_{02} \\ h_{10} & 1+h_{11} & h_{12} \\ h_{20} & h_{21} & 1 \end{bmatrix}$	$(h_{00}, h_{01}, \ldots, h_{21})$

Estymacja przekształcenia: metoda najmniejszych kwadratów

Jak w przypadku wielu modeli parametrycznych w teorii do problemu estymacji homografii między dwoma obrazami można użyć metody najmniejszych kwadratów, tzn. minimalizować wyrażenie:

$$E_{LS} = \sum_i ||r_i||^2 = \sum_i ||f(x_i; p) - x_i'||^2$$

gdzie $x_i' = f(x_i; p)$ jest interesującym nas przekształceniem, p jego parametrami.

Estymacja przekształcenia: ważona metoda najmniejszych kwadratów

Metoda najmniejszych kwadratów traktuje wszystkie dopasowania równorzędnie, tymczasem w praktyce w przypadku rozpoznawania obiektów dopasowania niezawsze będą dokładne, czasami mogą być nawet niepoprawne. Jednym z podejść by zniwelować wpływ słabych dopasowań jest nadanie obserwacjom wag ńiepewności". Minimalizowane wyrażenie jest wtedy postaci:

$$E_{WLS} = \sum_i \sigma_i^{-2} ||r_i||^2$$

Estymacja przekształcenia: *robust least squares*

Kolejnym podejściem do zminimalizowania wpływu outlierów i kiepskich dopasowań jest *robust least squares* - modyfikacja metody najmniejszych kwadratów. Minimalizowane wyrażenie jest wtedy postaci:

$$E_{RLS} = \sum_i \rho(||r_i||)$$

gdzie ρ jest funkcją kary. *Robust least squares* z powodzeniem eliminuje wpływ outlierów, ale nie jest rozwiązaniem idealnym - w przypadku, gdy outlierów jest dużo występują problemy ze zbieżnością (przy przekształceniach nieliniowych, gdy wymagane jest użycie metod iteracyjnych do znalezienia minimum). Metoda w tym zastosowaniu jest wypierana przez RANSAC i Least Median of Squares (LMS)

RANSAC: kroki

1. Wybieramy w sposób losowy minimalną liczbę punktów m niezbędnych do wyznaczenia parametrów modelu
2. Estymujemy parametry modelu na podstawie m wybranych obserwacji
3. Sprawdzamy, ile punktów z całej próby odpowiada wysestymowanemu modelowi (z założoną tolerancją d).
4. Jeżeli stosunek liczby punktów odpowiadających modelowi (tzw. inlierów) do liczby wszystkich punktów w próbie przekracza założony próg T, estymujemy model ponownie z użyciem tzw. zbioru konsensusu: m punktów użytych do wyestymowania modelu oraz inlierów.
5. W przeciwnym razie powtarzamy kroki 1-4

RANSAC

RANSAC czyli RANdom SAmple Consensus

Źródła

- R. Szeliski *Computer Vision: Algorithms and Applications*
- W szczególności: rozdział 2 (*Image formation*) i rozdział 6 (*Feature-based alignment*)