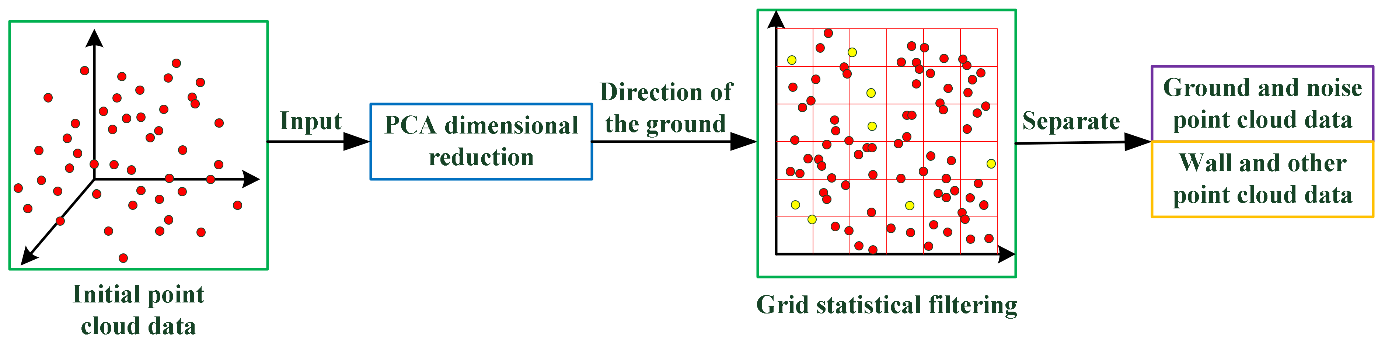
PCA = algoritmus slouží k redukci dimenze dat

https://www.keboola.com/blog/pca-machine-learning

Redukce šumu

Použití u Lidaru – redukce dat z 3D na 2D

PCA can be used to obtain the normal vector of a point cloud plane. In order to reduce the complexity of 3D point cloud processing and to effectively remove noise, we firstly project a 3D point cloud onto a two-dimensional plane



Firstly, dimensionality reduction via PCA is carried out for PCD. Assuming the input point cloud data S∈Rn×3, the covariance matrix Γ of S is performed a singular value decomposition (SVD) decomposition, therefore the corresponding eigenvectors. The first two eigenvectors are selected as dimensionality reduction matrix for the input S as follows:

Sq=SΔT

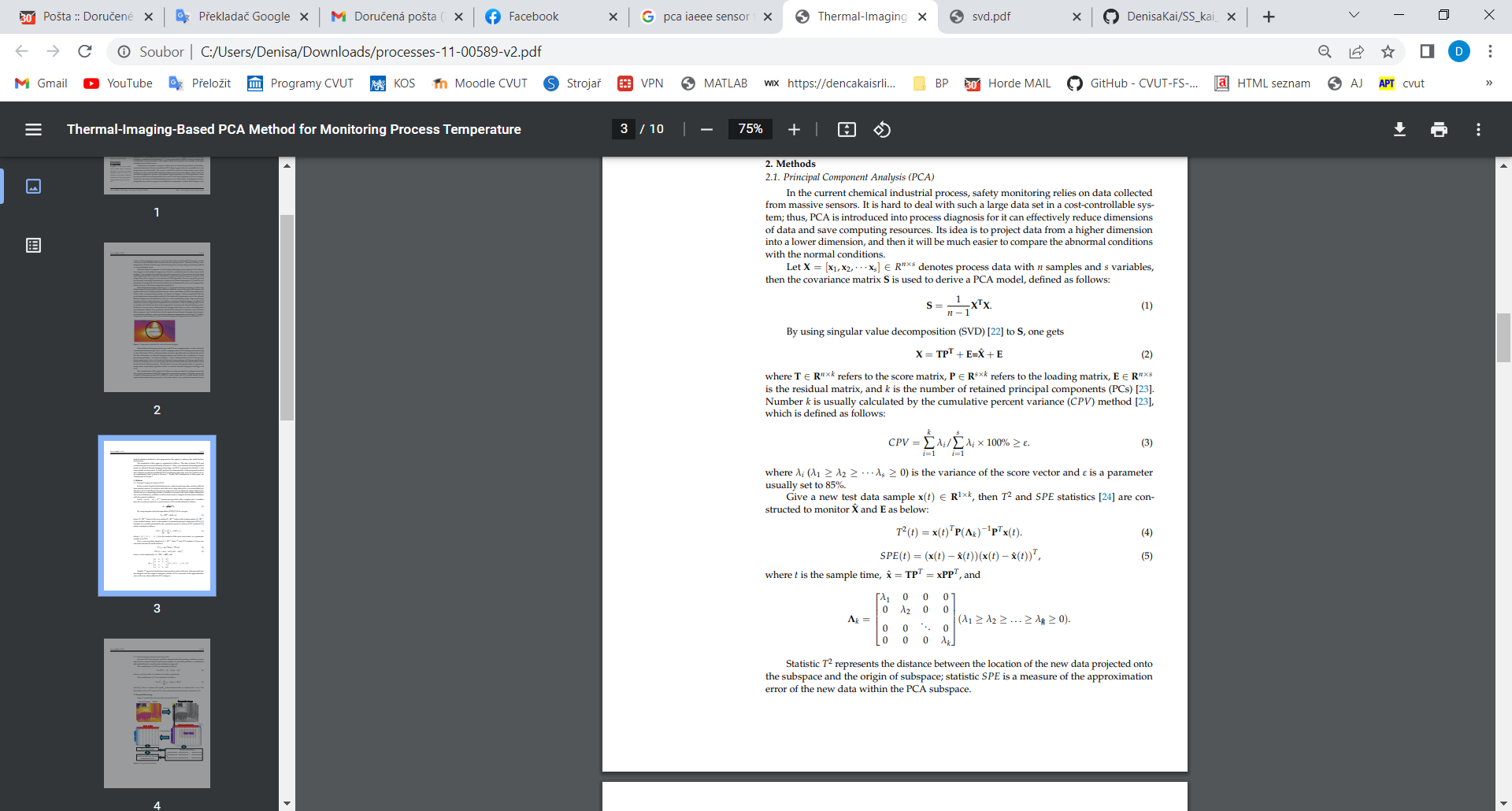
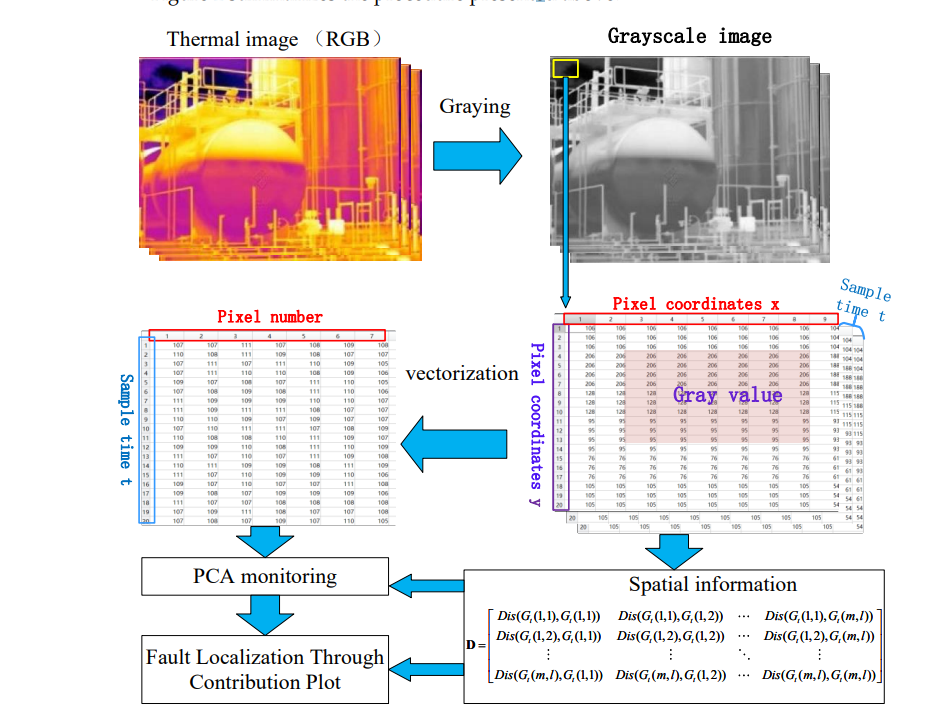
where ΔT∈R3×2 transforms the 3D point set S into a 2D point set denoted by Sq={xi,yi},i=1,…,n, which represents the point set after dimensional reduction.

Body 1-11 v článku

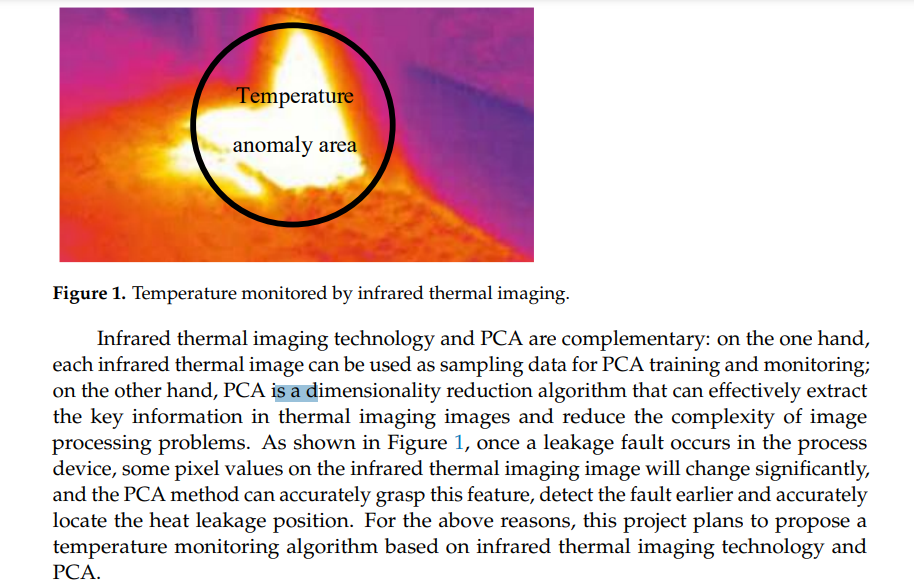
SVD = http://phoenix.inf.upol.cz/~konecnja/vyuka/2013/ALS1files/svd.pdf

Source: <https://www.mdpi.com/1424-8220/21/11/3703>

2.1. Principal Component Analysis (PCA) In the current chemical industrial process, safety monitoring relies on data collected from massive sensors. It is hard to deal with such a large data set in a cost-controllable system; thus, PCA is introduced into process diagnosis for it can effectively reduce dimensions of data and save computing resources. Its idea is to project data from a higher dimension into a lower dimension, and then it will be much easier to compare the abnormal conditions with the normal conditions. Let X = [x1, x2, · · · xs ] ∈ R n×s denotes process data with n samples and s variables, then the covariance matrix S is used to derive a PCA model, defined as follows: S = 1 n − 1 X TX. (1) By using singular value decomposition (SVD) [22] to S, one gets X = TPT + E=Xˆ + E (2) where T ∈ R n×k refers to the score matrix, P ∈ R s×k refers to the loading matrix, E ∈ R n×s is the residual matrix, and k is the number of retained principal components (PCs) [23]. Number k is usually calculated by the cumulative percent variance (CPV) method [23], which is defined as follows: CPV = k ∑ i=1 λi/ s ∑ i=1 λi × 100% ≥ ε. (3) where λi (λ1 ≥ λ2 ≥ · · · λs ≥ 0) is the variance of the score vector and ε is a parameter usually set to 85%. Give a new test data sample x(t) ∈ R 1×k , then T 2 and SPE statistics [24] are constructed to monitor Xˆ and E as below: T 2 (t) = x(t) T P(Λk ) −1 P T x(t). (4) SPE(t) = (x(t) − xˆ(t))(x(t) − xˆ(t)) T , (5) where t is the sample time, xˆ = TPT = xPPT , and Λk =      λ1 0 0 0 0 λ2 0 0 0 0 . . . 0 0 0 0 λk      (λ1 ≥ λ2 ≥ . . . ≥ λk ≥ 0). Statistic T 2 represents the distance between the location of the new data projected onto the subspace and the origin of subspace; statistic SPE is a measure of the approximation error of the new data within the PCA subspace.

Utřídění barev šedi pro termální kamery



V matlabu je implementovana funkce - dokumentace

https://www.mathworks.com/help/stats/pca.html