Feature detectors with rotation and scale invariance in CMOS-3D technologies for low power vision systems

Manuel Suárez Cambre Víctor M. Brea

Centro Singular de Investigación en Tecnoloxías da Información Universidade de Santiago de Compostela



Centro Singular de Investigación en **Tecnoloxías** da Información



Overview

Introduction

Thesis goal Methodology Context Hypothesis Focal Plane

Architecture for feature extractor Scale-Space on the Focal-Plane 3D-CMOS Error Sources SIFT Application

3 Proof of concept

Architecture Arrangement for implementation on a 2D technology Current and future work

Conclusions

Publications, projects and collaborations

Overview

- Introduction
- Thesis goal Methodology Context Hypothesis Focal Plane
- Architecture for feature extractor
 Scale-Space on the Focal-Plane 3D-CMOS
 Error Sources
 SIFT Application
- 3 Proof of concept

Architecture Arrangement for implementation on a 2D technology Current and future work

- 4 Conclusions
 - Publications, projects and collaborations





Implementation of the lowest level processing stage of the SIFT (Scale Invariant Feature Transform) algorithm on CMOS-3D technologies for high speed computation and low power consumption with image acquisition.



Methodology

Methodology

- Read literature
- $\,\triangleright\,$ Identify the context and the problem
- Identify contributions and set a goal
- Compare algorithms and select the best option
- Identify feasible simplifications
- Study the effects on performace
- Make a choice
- Circuit translation
 - Define a topology
 - Simulate to viability
 - Feedback
 - Final topology
 - Layout
 - Testing



Context

Applications

- ▷ Object detection & recognition
 - Surveillance
 - Quality control
- \triangleright Tracking
 - Surveillance
 - Robot Navigation
- ▷ Image algorithms
 - Stereo-Cameras calibration
 - Panoramic composition









Context

Applications

- ▷ Object detection & recognition
 - Surveillance
 - Quality control
- ▷ Tracking
 - Surveillance
 - Robot Navigation
- Image algorithms
 - Stereo-Cameras calibration
 - Panoramic composition





of concept Con

Context

Applications

- > Object detection & recognition
 - Surveillance
 - Quality control
- ▷ Tracking
 - Surveillance
 - Robot Navigation
- ▷ Image algorithms
 - Stereo-Cameras calibration
 - Panoramic composition

Bottom Left Camera (166 Features)



Bottom Right Camera (189 Features)





Context

Applications

- ▷ Object detection & recognition
 - Surveillance
 - Quality control
- ▷ Tracking
 - Surveillance
 - Robot Navigation
- \triangleright Image algorithms
 - Stereo-Cameras calibration
 - Panoramic composition





Hypothesis

The Problem

Feature repeatibility against scale and rotation changes, as well as partial occlusions or affine transformations with low power and real time

Implementations

- Few implementation
 - General purpose
 - GPUs
 - High power requirements
 - Around or slower than real time

Hypothesis

It is possible to embed one Feature extractor algorithm taking advantage of the benefits provided by CMOS-3D technologies



Conventional approach

- ▷ Two chips: acquisition + processing
 - Higher resolution
 - High memory and bandwidth requirements

Focal plane approach CMOS-2D

Focal plane approach CMOS-3D



Focal Plane

Conventional approach

Focal plane approach CMOS-2D

▷ One chip: acquisition + Processing

- SIMD
- High parallelism (High speed computation)
- Lower resolution
- Low-level processing

Focal plane approach CMOS-3D



Conventional approach

Focal plane approach CMOS-2D

Focal plane approach CMOS-3D

- one chip*: Acquisition + Processing
 - Same advantages as CMOS-2D approach
 - More parallelism
 - Less power consumption per memory access
 - Low, intermediate and high-level image processing



Image feature detectors: SIFT

Algorithms

- Low Accuracy/Low computation time
 - Harris
- ▷ High Accuracy/High computation time
 - Scale Invariant Feature Transform (SIFT)
 - Harris Affine
 - Hessian Affine
 - Speeded Up Robust Feature (SURF)



Image Feature Detectors

Algorithm major stages

- Gaussian Pyramid Generation (90 % of operations)
- Feature points location (1% of pixels in the image)
- Orientation assingment
- Descriptor vector

Hardware challenge

D To embed all these functions on a focal plane array with a small area occupation and a large resolution







SIFT analysis

SIFT analysis

- > The number of Octaves
- $\,\triangleright\,$ The number of Scales per Octave
- ▷ Thresholds (min/max or matching)



8/30

Introduction Architecture for feature extractor Proof of concept Conclusions Publications, projects and collaborations
SIFT analysis

SIFT analysis

- The number of Octaves
- $\,\triangleright\,$ The number of Scales per Octave
- ▷ Thresholds (min/max or matching)





The choice

What in analog? what in digital?

- \triangleright Analog
 - Advantages
 - > Gaussian filtering expensive in digital
 - > Analog RC network natural solution of Gaussian filtering
 - > Fully parallel
 - Drawbacks
 - > Other operations require area and long-term storage is an issue
 - > Mismatch
- Digital
 - All processing following the Gaussian filtering



Error Sources

Mismatch

- Manufacturing Processes
- \triangleright Variation from pixel to pixel
- ▷ Variation chip to chip
- Variation with respect to nominal values
 - Offsets
 - Gain
 - etc





Thesis goal Methodology Context Hypothesis Focal Plane

Architecture for feature extractor Scale-Space on the Focal-Plane 3D-CMOS Error Sources SIFT Application

3 Proof of concept

Architecture Arrangement for implementation on a 2D technology Current and future work

4 Conclusions

Publications, projects and collaborations



Architecture for feature extractor Proof of concept Conclusions Publications, projects and collaborations

Scale-Space on the Focal-Plane 3D-CMOS

Hardware Distribution by Layer

- \triangleright Tier 1 (Analog Domain)
 - QVGA array (320x240)
 - Processor (160x120):
 - > 4 Photodiodes
 - > Acquisition
 - > Analog Memories
 - > Singe slope A/D converter
 - Gaussian pyramid network
- \triangleright Tier 2 (Digital Domain)
 - 6 Registers per processor
 - A/D conversion
 - Derivatives
 - Extrema detection
 - Difference of Gaussians
- 1Gb DRAM Memory \triangleright
- Coprocessor





Processing Element or Cell





Proof of conc

ncept Conclusi

CMOS-3D Stack Architecture

Acquisition

One acquisition block for four photosensors

$$P_{P_{Si}} = V_{ref} + \frac{C}{C_{Pi}} [V_S(t_0) - V_S(t_1)] - V_Q$$

$$\triangleright \quad V_{out} = V_{ref} + \frac{C}{C_{Pi}} [V_S(t_0) - V_S(t_1)]$$

 \triangleright CDS

$$\triangleright C = C_{Si} = 200 fF$$

Conversion





CMOS-3D Stack Architecture

Acquisition

Conversion

- ▷ Single Slope In-pixel A/D Converter distributed through two tiers
- \triangleright One A/D converter for four photosensors

$$\triangleright V_{out} = -K(V_{ramp} - V_{Si}) + V_Q$$

- 8 bit resolution \triangleright
- 100us per conversion \triangleright





CMOS-3D Stack Architecture

Gaussian Filtering

 $\,\triangleright\,\,$ Gaussian: best function points detection

-
$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y}{2\sigma^2}}$$

Scale = Gaussian convolution with input image

-
$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$

Solution: RC network

$$\triangleright \sigma = \sqrt{\frac{2t}{RC}}$$

- Our Implementation
 - Switched-Capacitors network





Introduction

Proof of conce

ncept Conclus

ns Publications, projects and collaboration

CMOS-3D Stack Architecture

Gaussian Filtering

▷ 2D Network

-
$$V_{ij}(n) = V_{ij}(n-1) + [V_{i+1j}(n-1) + V_{i-1j}(n-1) + V_{ij+1}(n-1) + V_{ij-1}(n-1) + V_{ij}(n-1)] \frac{C_E}{C_{P_i}}$$

1) - 4 $V_{ij}(n-1)$] $\frac{C_E}{1+4\frac{C_E}{C_{P_i}}}$

Software Kernel

-
$$V_{ij}(n) = V_{ij}(n-1) + [V_{i+1j}(n-1) + V_{i-1j}(n-1) + V_{ij+1}(n-1) + V_{ij-1}(n-1) + V_{ij-1}(n-1) - 4V_{ij}(n-1)] \frac{e^{-\frac{1}{2\sigma^2}}}{1+4e^{-\frac{1}{2\sigma^2}}}$$

- $\triangleright \ \sigma_0 = (2ln \frac{C}{C_E})^{-1/2}$
- \triangleright With the clock cycles:

-
$$\sigma(n) = (\frac{2nC_E}{4C_E + C_{Pi}})^{1/2}$$





Architecture for feature extractor Proof of concept Conclusions Publications, projects and collaborations

CMOS-3D Stack Architecture





Error Sources

Error Sources

Mismatch $C - C_{Si}$ \triangleright

-
$$V_{out} = V_{ref} + \frac{C}{C_{Si}} [V_S(t_0) - V_S(t_1)]$$

Mismatch $C_{Si} - C_F$ \triangleright

-
$$V_{ij}(n) = V_{ij}(n-1) + [V_{i+1j}(n-1) + V_{i-1j}(n-1) + V_{ij+1}(n-1) + V_{ij-1}(n-1) - 4V_{ij}(n-1)] \frac{\frac{C_E}{C}}{1 + 4\frac{C_E}{C}}$$

Charge Injection and Feedthrough \triangleright





Proof of conce

cept Conclusio

Publications, projects and collaborations

SIFT Application: Object Detection

Software vs Mismatch

- ▷ The variation of local *sigma*
- Some points can become extrema
- The same point in a tranformed image suffers different filtering
- Incorrect matches increase
- Accuracy degradation

Analysis Conditions







SIFT Application: Object Detection

Software vs Mismatch

Analysis Conditions

- Object detection with the SIFT algorithm
- Gaussian Network implemented in Matlab
- ▷ Image 320x240
- $\triangleright C = 200 fF, C_E = 20 fF$
- $\triangleright \ \ Capacitances \ \ Variations \\ 6\sigma = \sqrt{C}$

$$\triangleright \ \sigma_{\rm SIFT} \to \sigma = \sigma(n) \to n$$





Overview

- Introduction
- Thesis goal Methodology Context Hypothesis Focal Plane
- Architecture for feature extractor
 Scale-Space on the Focal-Plane 3D-CMOS
 Error Sources
 SIFT Application
- 3 Proof of concept

Architecture Arrangement for implementation on a 2D technology Current and future work

Conclusions

Publications, projects and collaborations



2D rearrangement

Some numbers

- D 180nm technology
- ▷ Tiers 1 & 2 merged in one
- ▷ 2 registers per cell
- \triangleright 176x120 pixels in a 5x5mm²

- \triangleright cell 44x44 μ m²
- $\triangleright \sigma_0 = 0,48$
- ▷ 75 nW/pixel at 30 frames/s





Proof of concep

cept Conclusion

Publications, projects and collaborations

Results: circuits simulation 16x16 array





The baby





Current and future work

SETUP for test

- PCB for chip fastening and iteration
- Control
 - Signal pattern generator
 - FPGA
- Stand-alone system





Overview

Architecture Arrangement for implementation on a 2D technology

Conclusions



- $\,\triangleright\,$ An Architecture for Gaussian pyramid generation was proposed
- ▷ For implementation in 130nm Tezzaron 3D-Technology
- A switched diffussion network
- $\triangleright \,\, \sigma$ of filtering controled by clk cycles
- ▷ 3 Octaves
- Scales programmable by user
- For feature extractor system
- ▷ Proof of concept in a 2D technology
- Array 176x120
- Gaussian pyramid
- ▷ Test is coming



Overview

- Introduction
- Thesis goal Methodology Context Hypothesis Focal Plane
- Architecture for feature extractor Scale-Space on the Focal-Plane 3D-CMOS Error Sources SIFT Application

3 Proof of concept

Architecture Arrangement for implementation on a 2D technology Current and future work

Conclusions

5 Publications, projects and collaborations



Publications

Publications

2 journals + 11 conferences

Journal

- A hierarchical vision processing architecture oriented to 3D integration of smart \triangleright camera chips. Ricardo Carmona-Galán, ákos Zarándy, Csaba Rekeczky, Péter Földesy, Alberto Rodríguez-Pérez, Carlos Domínguez-Matas, Jorge Fernández-Berni, Gustavo Liñán-Cembrano, Belén Pérez-Verdú, Zoltán Kárász, Manuel Suárez-Cambre, Victor Brea-Sánchez, Tamás Roska, Ángel Rodríguez-Vázquez. J. Syst. Architect. (2013), http://dx.doi.org/10.1016/j.sysarc.2013.03.002
- \triangleright CMOS-3D Smart Imager Architectures for Feature Detection. Suarez, M.; Brea, V.M.; Fernandez-Berni, J.; Carmona-Galán, R.; Liñán, G.; Cabello, D.; Rodríguez-Vázquez, A., Emerging and Selected Topics in Circuits and Systems, IEEE Journal (JETCAS) on , vol.2, no.4, pp.723,736, Dec. 2012.



Publications

Conference

- A 176x120 Pixel CMOS Vision Chip for Gaussian Filtering with Massivelly Parallel \triangleright CDS and A/D-Conversion. Suárez, M.; Brea, V.M.; Fernández-Berni, J.; Carmona-Galán, R.; Cabello, D.; Rodríguez-Vázquez, A., 21th European Conference on Circuits Theory and Design (ECCTD). Sep. 2013. (En proceso de aceptación)
- FPGA-oriented, fast and efficient calculation of orientation for SIFT keypoints. Illade-Quinteiro J.; Brea, V.M.; Suárez, M.; Carmona-Galán, R.; Rodríguez-Vázquez, A. (En proceso de aceptación).
- \triangleright In-pixel generation of gaussian pyramid images by block reusing in 3D-CMOS. Suárez, M.; Brea, V.M.; Cabello, D.; Carmona-Galán, R.; Rodríguez-Vázquez, A., Circuits and Systems (ISCAS), 2012 IEEE International Symposium on , vol., no., pp.2649,2652, 20-23 May 2012.
- \triangleright Scale- and rotation- invariant feature detectors on Cellular Processor Arrays. Fernández, N.A.; Brea, V.M.; Suaárez, M.; Cabello, D., Circuits and Systems (ISCAS), 2012 IEEE International Symposium on , vol., no., pp.2657,2660, 20-23 May 2012.



Publications

Conference

- Evidence of the lateral collection significance in small CMOS photodiodes. \triangleright Blanco-Filgueira, B.; Lopez, P.; Doge, J.; Suárez, M.; Roldan, J.B., Circuits and Systems (ISCAS), 2012 IEEE International Symposium on , vol., no., pp.3098,3101, 20-23 May 2012
- \triangleright A CMOS-3D Reconfigurable Architecture with In-pixel Processing for Feature Detectors. Suárez, M.; Brea, V.M.; Pardo, F.; Carmona-Galán, R.; Rodríguez-Vázquez, A., 3D Systems Integration Conference (3DIC), 2011 IEEE International, vol., no., pp.1,8, Jan. 31 2012-Feb. 2 2012.
- \triangleright Switched-Capacitor Networks for Scale Space Generation. M. Suárez, V.M. Brea, D. Cabello F. Pozas-Flores, R. Carmona-Galán and A. Rodríguez Vázquez. 20th European Conference on Circuits Theory and Design (ECCTD), pp 189-192. Linköping, Sweden. 29-31 August 2011.
- \triangleright In-Pixel ADC for a Vision Architecture on CMOS-3D Technology. M. Suárez, V.M. Brea, Carlos Domínguez Matas, Ricardo Carmona, Gustavo Liñán and ángel Rodríguez Vázquez. IEEE International 3D System Integration Conference. P23. Munich, Germany. 16-18, 2010.



Conference

- A 3D chip architecture for optical sensing and concurrent processing. ángel Rodríguez-Vázquez, Ricardo Carmona, Carlos Domínguez Matas, Manuel Suárez-Cambre, Victor Brea, Francisco Pozas, Gustavo Liñan, Peter. Foldessy, Akos. Zarandy and Csaba Rekeczky. 12-15 April 2010. Brusseles, Belgium. Proc. SPIE 7726, 772613 (2010); doi: doi:10.1117/12.855027.
- Offset-Compensated Comparator with Full-Imput Range in 150nm FDSOI
 CMOS-3D Technology. M. Suárez and V.M. Brea, Carlos Domínguez Matas, Ricardo
 Carmona, Gustavo Liñán and ángel Rodríguez Vázquez. pp. 184-187. Latin American
 Symposium on Circuits and Systems (LASCAS). Iguazu Falls, Brazil. 24-26 February 2010.
- Template-Oriented Hardware Design based on Shape Analisys of 2D CNN
 Operators in CNN Template Libraries and Applications. Natalia A. Fernández García,
 M. Suárez, V.M. Brea and D. Cabello. 11th International Workshop on Cellular Neural
 Networks and their Applications. pp. Santiago de Compostela, Spain. 14-16 July 2008.





- ▷ VISCUBE (ONR)
- ▷ Xunta de Galicia through project 10PXIB206037PR,
- ▷ MICINN through TEC2009-12686
- ▷ MINECO TEC2012-38921-C02
- ▷ CENIT ADAPTA



Technology transfer

- ▷ IMAGE PROCESSOR FOR FEATURE DETECTION
- ▷ PCT P201200090
- ▷ US 13/417,279

Thesis

▷ Expected Thesis defense: to the end of 2013



Collaborators











Thank you for your attention!!!

