Building Identification and Gaining Access Using AR Drone

By:
Smith Gupta 11720
Dhruv Kumar Yadav 11253

Contents

- What is A.R. Drone?

- Problem Statement

- Object Classification
  - Representation
  - Learning
  - Matching

- Dataset

- Future Work

- References

Source: appadvise.com

WHAT IS AR DRONE??

QUADCOPTER especially UAV

Source: quadcopters.co.uk

- **Aerial vehicle propelled by 4 rotors**

- **2 sets of identical fixed pitched propellers**; 2 clockwise (CW) and 2 counter-clockwise (CCW)

- Use **variation of RPM to control lift/torque**

- **Control of vehicle** motion is achieved by **altering the rotation rate of one or more rotor discs, thereby changing its torque load and thrust/lift characteristics**

Source: wikipedia.com

AR Drone

Source: jazarah.net

- AR Drone is widely used Unmanned Aerial Vehicle

- Heavily used as Research platform due to its robustness, mechanical simplicity, low weight and small size

- It has been used for object following, position stabilisation, autonomous navigation and has wide applications in military reconnaissance and surveillance, terrain mapping and disaster management
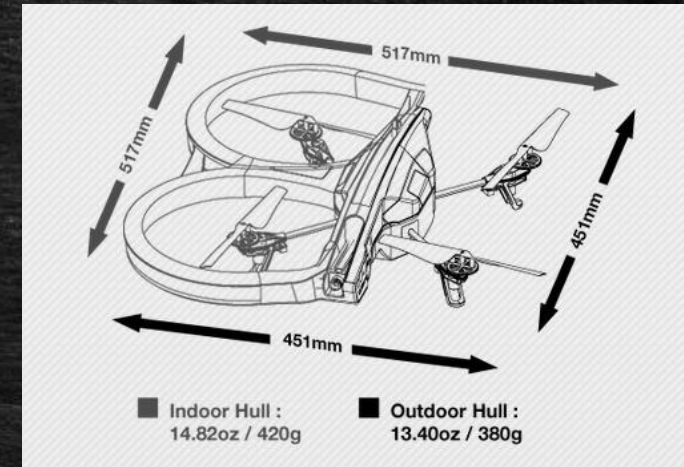
AR Drone PLATFORM

- HARDWARE
  - **WI-FI** 802.11 b/g/n to communicate via ad-hoc network
  - **1Gbit DDR2 RAM at 200MHz**
  - **Front Camera**
    - 720p, 30fps HD Video Recording
    - 75° × 60° field of view
  - **Bottom Camera**
    - 176 × 144 p, 60 fps vertical QVGA camera
    - Field of view 45° × 35°

Source: *ardrone2.parrot.com/ardrone-2/specifications/*

Source: ardrone2.parrot.com/ardrone-2/specifications/

AR Drone PLATFORM

- SOFTWARE
  - Linux 2.6.32
  - Communication via Wi-Fi AdHoc from ground server
  - **Smart phone Application** for Android and IOS based platform available

517mm

517mm

451mm

451mm

Indoor Hull :
14.82oz / 420g

Outdoor Hull :
13.40oz / 380g

iFixit

Source: *ardrone2.parrot.com/ardrone-2/specifications/*

PROBLEM STATEMENT

- **Autonomous identification of large structures like buildings from aerial imagery using an AR Drone**

- **The training and test data sets for identification include images of buildings at IIT Kanpur captured using front camera of AR Drone**

- The identification task is to be done at run-time i.e., during the flight of the drone

- Upon recognition the Quadcopter gain access in the building via open portal - window or door

MOTIVATION

- 4th mission of International Aerial Robotics Competition (IARC)
  - UAV flying 3 Km to an abandoned village and identifying a structure based on symbol on the building
  - Upon identification the UAV has to access the structure through open portals (doors, windows, other openings) that had to be identified by the UAV

- Going beyond identification we plan to move on to a bigger project
  - Developing an autonomous system that will help people navigate through an unknown/GPS denied environment
  - This can be used for finding routes, or even as a tour guide

APPROACH

Feature Detection and Description

↓

Bag of Visual Words Model

↓

Keypoint Classification using SVM

FEATURES
OR
INTEREST POINTS

Interesting points on the object that can be extracted to locate the object

Object should be detectable under change of scale, rotation and noise

Properties of Feature Points

- **Repeatable** – Feature in one image can be found in other image

- **Distinctive Description** – Each feature has distinctive property

- **Locally Salient** – Occupy small area of image; robust to clutter
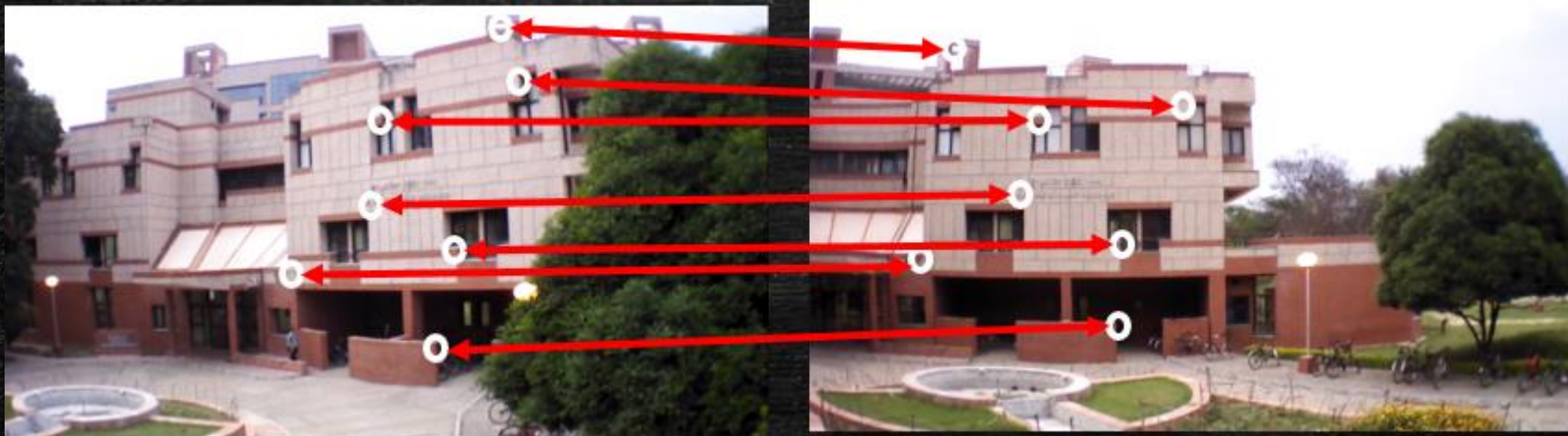
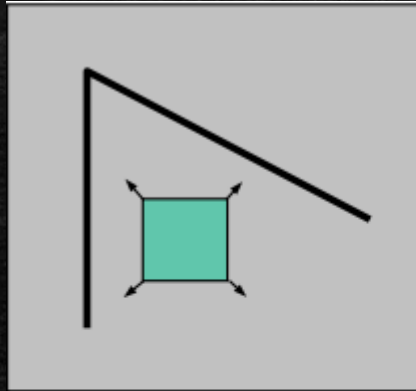DATABASE   TEST IMAGE

What are Features?

- Harris Corner Point

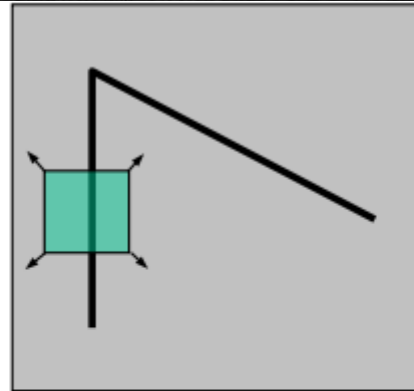- SIFT Detector

Feature Detection
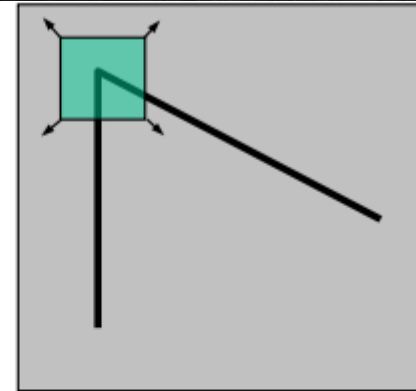# Harris Corner Point Detector

- Intuition: Match corners i.e. points with large variation in neighbourhood

"flat" region:
no change in all
directions

"edge":
no change along
the edge direction

"corner":
significant change
in all directions

Feature Detection
Harris Corner Point Detector

- Thus look at change in Intensity value

$$E(u,v) = \sum_{(x,y) \in W} [I(x+u, y+v) - I(x,y)]^2$$

- Using Taylor series expansion

$$I(x+u, y+v) \approx I(x,y) + \frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v \quad \approx I(x,y) + [I_x \ I_y] \begin{bmatrix} u \\ v \end{bmatrix}$$

- Thus

$$E(u,v) \approx \sum_{(x,y) \in W} \left[[I_x \ I_y] \begin{bmatrix} u \\ v \end{bmatrix} \right]^2$$

$$E(u,v) = \sum_{(x,y) \in W} [u \ v] \underbrace{\begin{bmatrix} I_x^2 & I_x I_y \\ I_y I_x & I_y^2 \end{bmatrix}}_{H} \begin{bmatrix} u \\ v \end{bmatrix}$$

Source: slides by Steve Seitz, Kristen Grauman, DevaRamanan

Feature Detection
Harris Corner Point Detector

- Thus find the direction ([u v]) will result in largest and smallest Eigen values

- We can find the value by looking at Eigen vectors of H

- For pixel/ patch to be corner point; even the smallest Eigen value should be large enough

- Apart from smallest Eigen vector we can also look at Harris Corner Point

$$f = \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2} = \frac{determinant(H)}{trace(H)}$$

Source: slides by Steve Seitz, Kristen Grauman, DevaRamanan

DRAWBACKS
Harris Corner Point Detector

REFERENCE: *A COMBINED CORNER AND EDGE DETECTOR by Chris Harris & Mike Stephens [1988]*

- Harris Corner point is very sensitive to changes in image scale

- Although Harris corner can detect corners and highly textured points, it is not a good feature for matching images under different scales

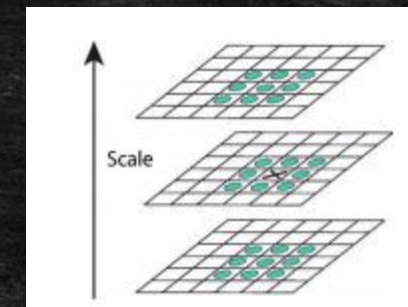## NOT DESIRABLE FOR OUR PROBLEM
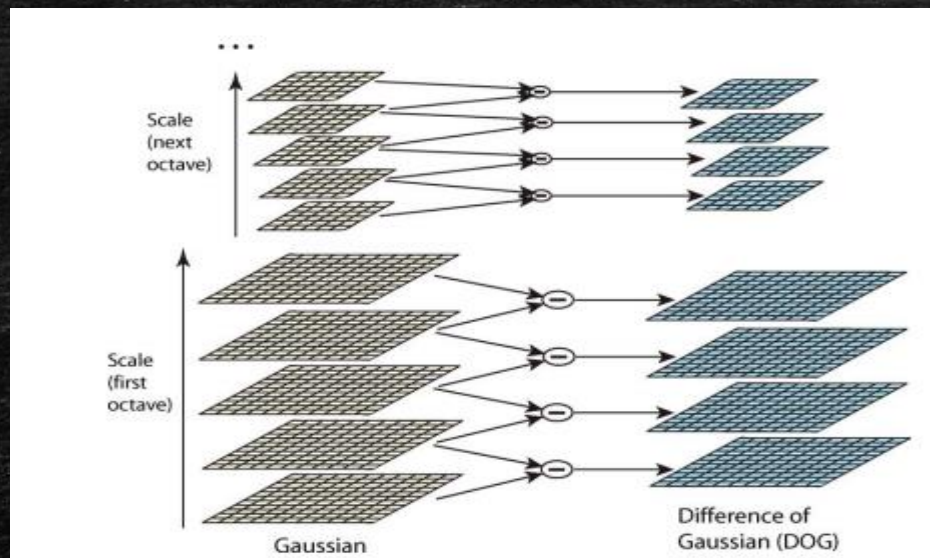
Feature Detection and Description
SIFT: Scale Invariant Feature Transform

- Intuition: Construct Scale space. Find the interest point in DoG space (Difference of Gaussian) by comparing a pixel with its 26 neighbouring pixels in the current and adjacent scales

$$\sigma \nabla^2 G = \frac{\partial G}{\partial \sigma} \approx \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k\sigma - \sigma}$$

$$G(x, y, k\sigma) - G(x, y, \sigma) \approx (k - 1)\sigma^2 \nabla^2 G.$$

- Eliminate edge points by constructing H matrix and look for Harris function with non maximal suppression
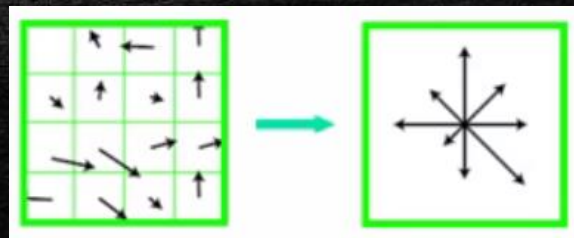
Feature Detection and Description
SIFT: Scale Invariant Feature Transform

- Orientation Assignment to each keypoint for rotation invariant. Descriptor now is represented relative to this orientation

- For each image sample L(x,y) at specific scale, gradient magnitude and orientation on the Gaussian smoothed images is pre-computed

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}$$
$$\theta(x, y) = \tan^{-1}((L(x, y+1) - L(x, y-1))/(L(x+1, y) - L(x-1, y)))$$

- Create a weighted direction histogram in neighborhood of keypoint 36 bins

- Peak in the histogram correspond to the orientations of the patch

Feature Detection and Description
SIFT: Scale Invariant Feature Transform

REFERENCE:*Distinctive Image Features from Scale-Invariant Keypoints by David Lowe [2004]*

- FEATURE DESCRIPTOR

- Based on 16*16 patches

- 4*4 subregions

- 8 bins in each subregion

- 4*4*8=128 dimensions in total

Image gradients → Keypoint descriptor
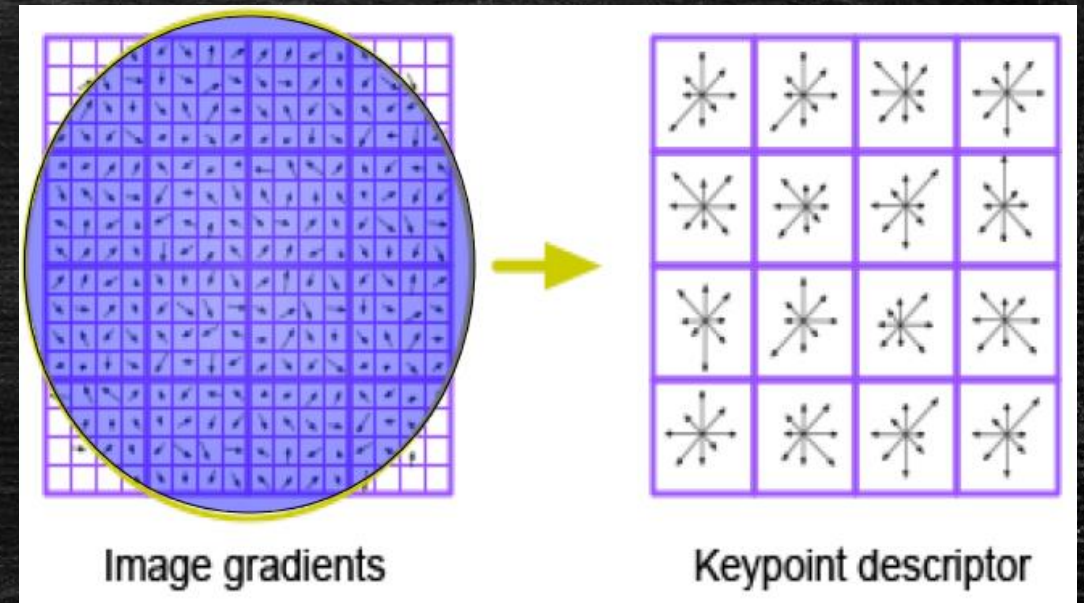
Source: Jonas Hurreimann

BAG OF VISUAL WORDS

- Split space of feature descriptors into multiple clusters using k-means algorithm

- Each resulting cluster cell is then mapped to a visual word

- Each Image is represented as histogram of these visual words

Vector quantization
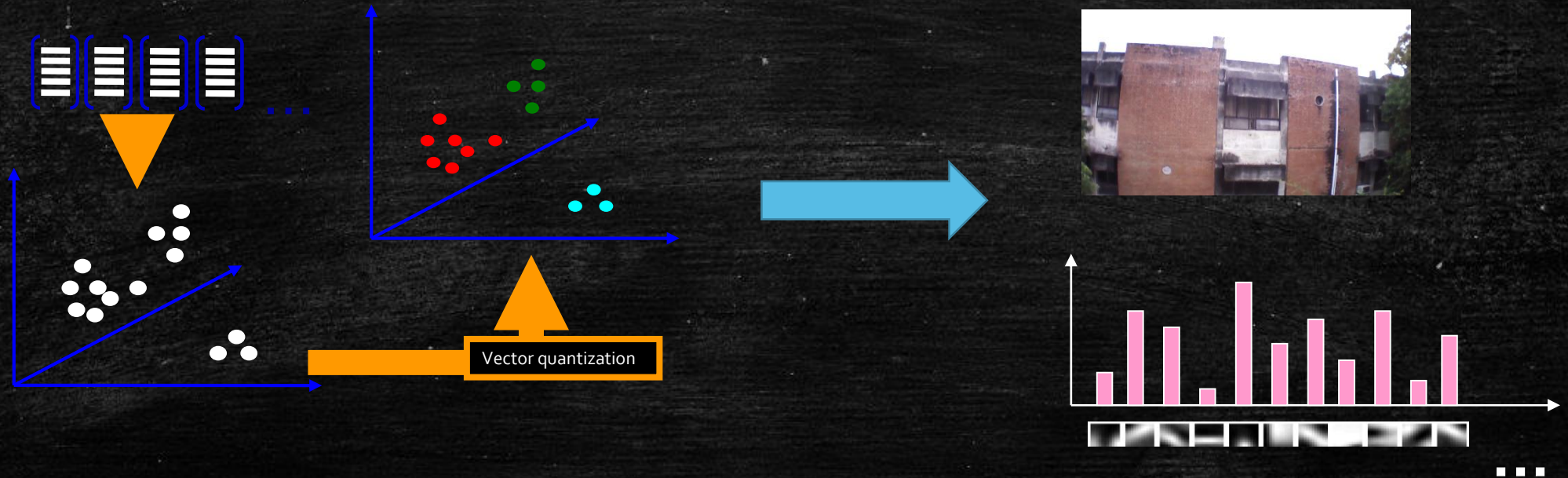
LEARNING: Support Vector Machine

- Types of Classifiers:
  - One vs. All
  - One vs. One

- One vs all deals with all the data of all the samples thus consume more time

- Thus we plan to use **one vs one** to increase classification speed

Dataset

Images of buildings at IIT Kanpur

Tools Used

AR Drone SDK 1.8

ffmpeg libraries

Working with Test Data

- For test data we use video streaming captured by the front camera of the drone

- Codec used by AR Drone 2.0 is H.264/MPEG-4

- Images can be extracted from this streaming video by using ffmpeg libraries

- These images would be given as inputs to SIFT algorithm at a certain frequency

- Each test image is matched with the database and probability measure is assigned

- If greater than threshold; image is successfully classified else look for another measure
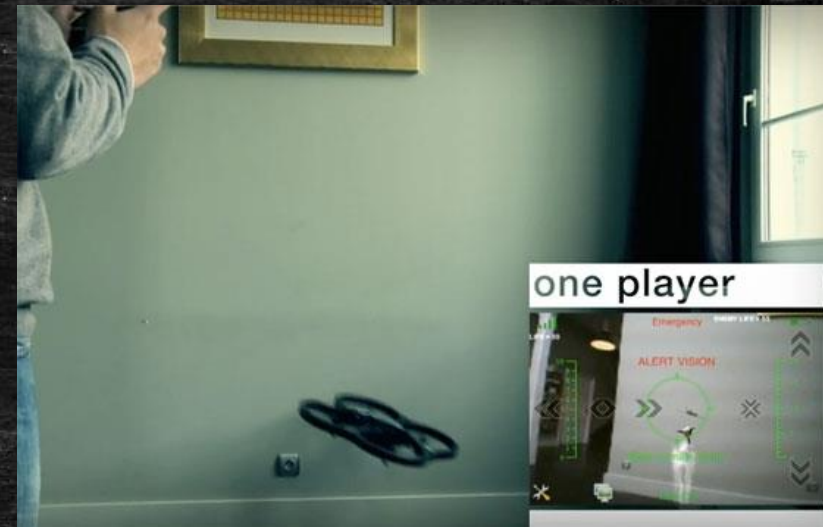
FUTURE WORK

- Compare between various Techniques like
  - vocabulary tree
  - bag or words with k nearest neighbour

- Build SKYCALL like system at IIT KANPUR
  - An autonomous flying quadcopter and a personal tour guide build at MIT Senseable City Lab
  - Guide prompts the users for the destination they want to reach
  - A mobile based application is being developed through which a user can "call" the guide for assistance

SUMMARY

SIFT

Bag of Visual Words Model

SVM Classifier

Principle of Support Vector Machines (SVM)

Input Space   Feature Space
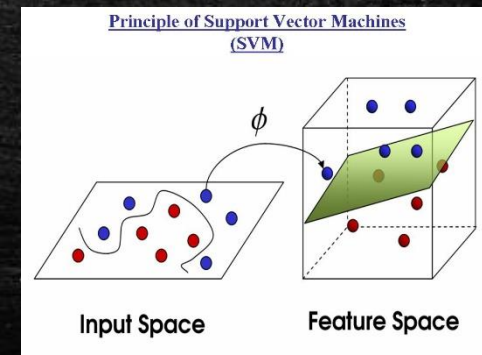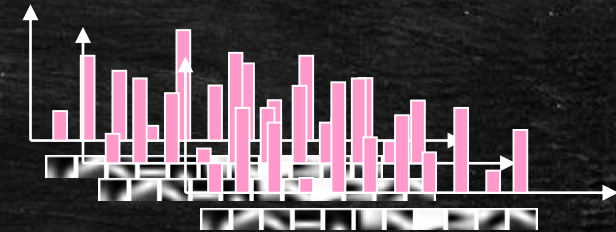
REFERENCES

- [1] *AR-Drone as a Platform for Robotic Research and Education*- Tomas Krajník, Vojtech Vonasek, Daniel Fiser, and Jan Faigl   {2011}

- [2] *Image target identification of UAV based on SIFT.* - Xi Chao-jian,Guo San-xue   {2011}

- [3] *Architectural Building Detection and Tracking in Video Sequences Taken by Unmanned Aircraft System (UAS)* Qiang He, Chee-Hung, Henry Chu and Aldo Camargo   {2013}

- [4] *Contextual Bag-of-Words for Visual Categorization -* Teng Li, Tao Mei, In-So Kweon   {2011}

- [5] *A SIFT-SVM METHOD FOR DETECTING CARS IN UAV IMAGES -* Thomas MORANDUZZO and Farid MELGANI   {2012}

- [6] *Multi-Information based Safe Area Step Selection Algorithm for UAV'S Emergency Forced Landing -* Aiying Lu, Wenrui Ding and Hongguang Li   {2013}

Q-A & SUGGESTIONS????

Source: *http://s1.reutersmedia.net/*