

Dynamic Active Search for Quick Object Detection with Pan-Tilt-Zoom Camera

Takahito KAWANISHI Hiroshi MURASE Shigeru TAKAGI Martin WERNER
NTT Communication Science Laboratories Email: kawanisi@eye.brl.ntt.co.jp

PURPOSE

Detect a known 3D object with a pan-tilt-zoom camera from a 3D scene accurately and quickly

Target

pan-tilt (wide view)

Zoom

Pan

Tilt

zoom-in (high resolution)

Detect!

PROBLEM

- Combination of pan-tilt-zoom parameter become large
- Many reference images are required for changes in illumination, poses, and sizes

Many reference images

Many camera operations

DETAILS

- Active Search with Union Histogram
 - A union histogram is defined by left equation.
 - $S(U, F)$ is the upper bounds of $S(R_m, F)$
 - A negative result of Active Search with Union histogram guarantees the absence of all the reference images.
 - In prediction stages, unnecessary zoom operations can be efficiently pruned.
- Parallel Active Search
 - Parallel Active Search prunes the mutual search spaces among multiple reference images' search.
 - The difference between reference histograms is calculated beforehand.
 - The upper bound of $S(R_n, F)$ can be calculated from $S(R_m, F)$ and the difference between R_m and R_n during Active Search with R_m .

$U^i = \max(R_i^1, R_i^2, \dots)$
 i is the color code

$S(U, F)$

$S(R_m, F) < S(U, F)$

$S(R_n, F) < S(R_m, F) + |R_n - S(R_m, R_n)|$

PREVIOUS WORK

Active Search on an image

Feature: color histogram
R: the histogram of reference images
F: the histogram of focus regions

Similarity:
histogram intersection
 $S(R, F)$: histogram intersection between R and F

Upper Bound of similarity for a neighboring focus region

$S(R, F_b) < S(R, F_a) + |F_b - F_a|$

$|F_b - F_a|$ denotes the number of pixels in F_b not in F_a

If upper bound of F_b is less than threshold, matching on F_b is pruned

SOLUTION

- Dynamic Camera Control
wide-angle-first and best-direction-first strategy
- Union histogram and Parallel Active Search
Active Search with a Union Histogram
Reference histograms for prediction are merged into a union histogram.
- Parallel Active Search
Reference histograms for verification prune search spaces of each other.

Prediction Stage
(No misses)

Zoom in

Verification Stage
(No misses and No redundant detection)

Few camera operations

RESULTS

Method	Camera Control Time	Active Search Time	Total Time
Simple Camera Control + Active Search	20.7s	23.9s	44.6s
Dynamic Camera Control + Parallel Active Search	6.0s	14.5s	20.5s
Dynamic Camera Control + Union Histograms	6.0s	3.6s	9.6s

Effect of Dynamic Camera Control

Effect of improved Active Search

Input Scene

Reference Image