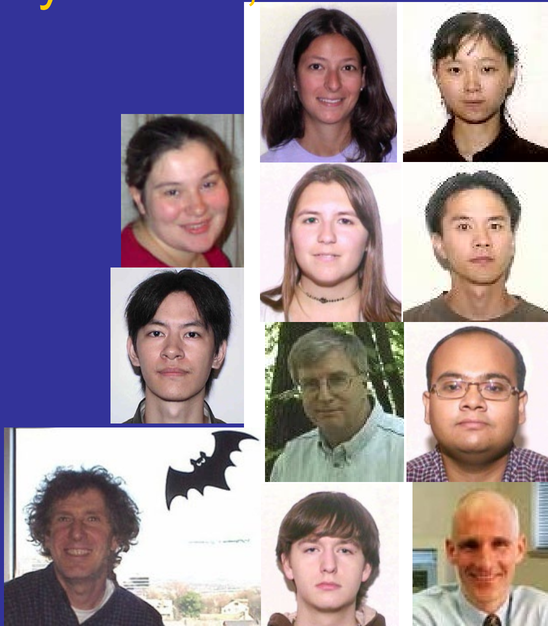
CS 585 Lecture on Single-View Multi-Object Tracking explained with Research on *Bat Behavior*

Margrit Betke
Computer Science Department
Boston University

February 2024

Acknowledgements

- NSF EIA-ITR 0326483: $2.4 Million (2004-2009) and 0910908 IIS $2.8 Million (2009-2016)
- Collaborators at Boston University:
  - Computer Science: Prof. Stan Sclaroff, Angshuman Bagchi, Stephen Crampton, Diane Hirsh, Eric Immermann, Edward Lee, Lisa Premerlani, Marianne Procopio, Shuang Tang, Zheng Wu
  - Biology: Prof. Thomas Kunz, Louise Allen, Nickolay Hristov, Jason Horn, Nathan Fuller, Jonathan Reichard
  - Geography: Prof. Cutler Cleveland
- Collaborators at University of Tennessee:
  - Biology: Prof. Gary McCracken
- Collaborators at MIT:
  - Mechanical Engineering: Prof. Nicholas Makris



Bat Colonies in Southwest

Colonies of Brazilian free-tailed bats may be the largest aggregations of mammals in the world

Why study bats?

- Colonies of bats represent some of the largest aggregations of mammals known to humankind
- Second most diverse order of mammals; 1/5th of mammalian species are bats
- Bats provide valuable services to humans
  - Insectivorous bats consume pest insects
  - Nectarivorous bats are pollinators
  - Frugivorous bats aid in seed dispersal
- Build our understanding of disease ecology
  - Bats are reservoirs for emerging diseases
  - North American bats are being devastated by a fungal disease

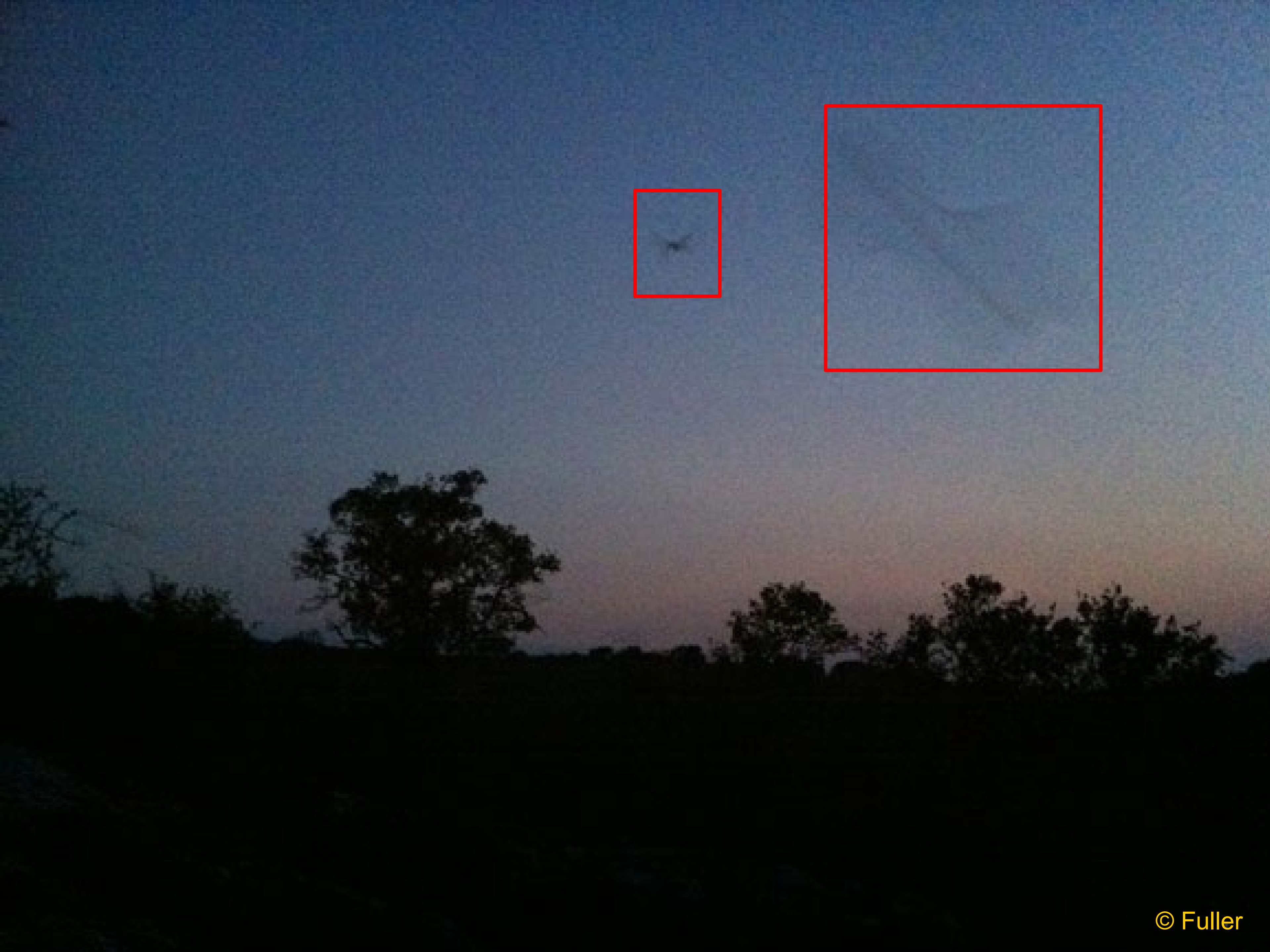
Corn earworm

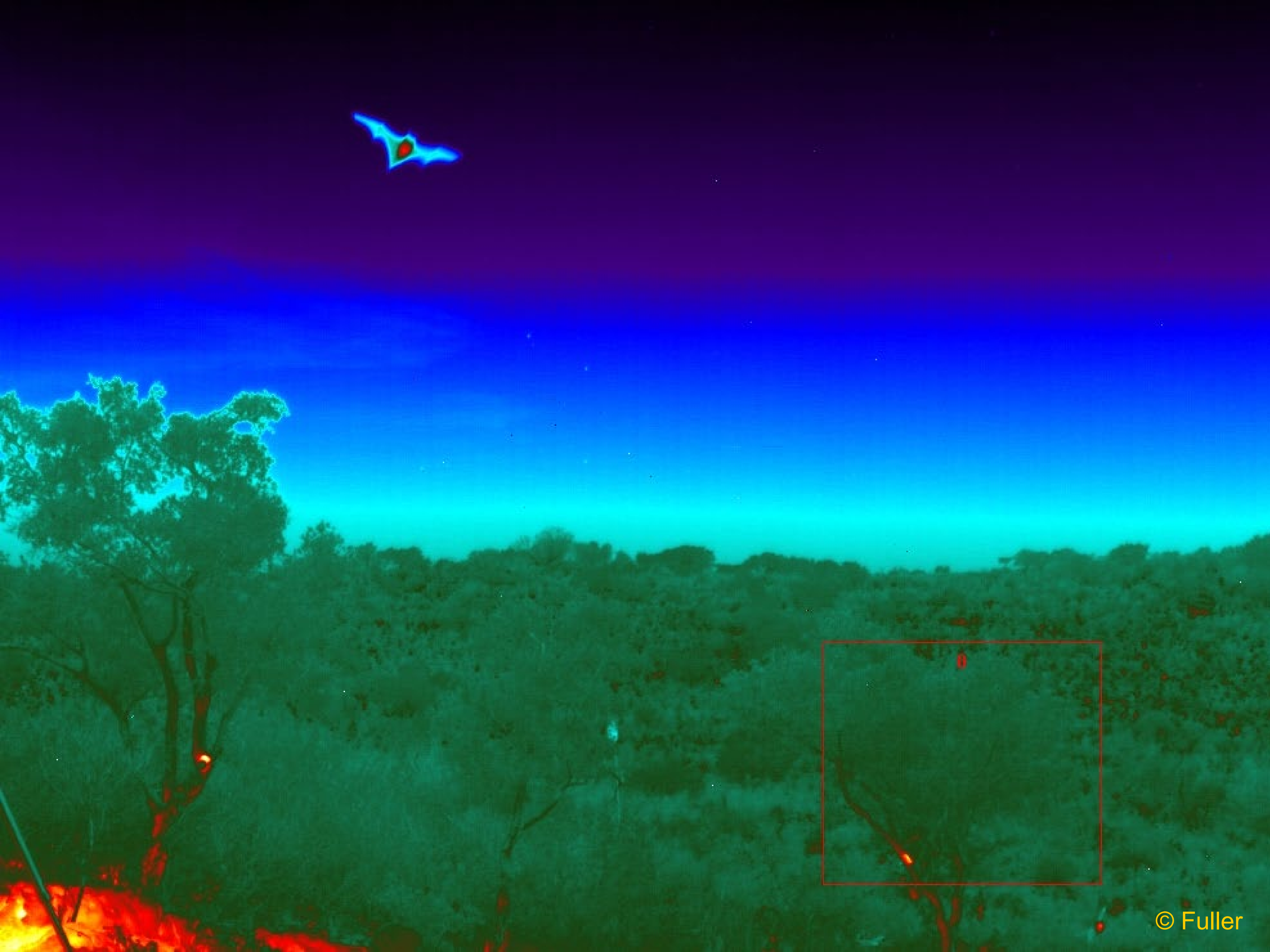

© Fuller


© Fuller

Challenges of studying bats

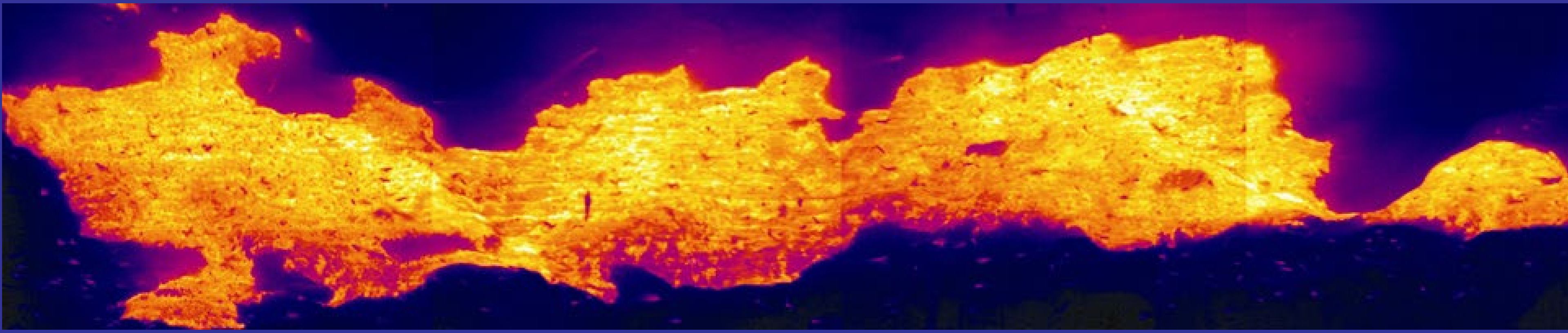
- Bats are nocturnal
- Bats fly quickly and are highly maneuverable
- Bats are sensitive to disturbance and learn quickly
  - Cue in to the sounds of trap setup
  - Have multiple exits from their roost
  - Constant trapping will lead to roost abandonment
- Recapture rates are very low (10% is considered average)
- Developing reliable and cost-effective methods of non-invasive monitoring is vital

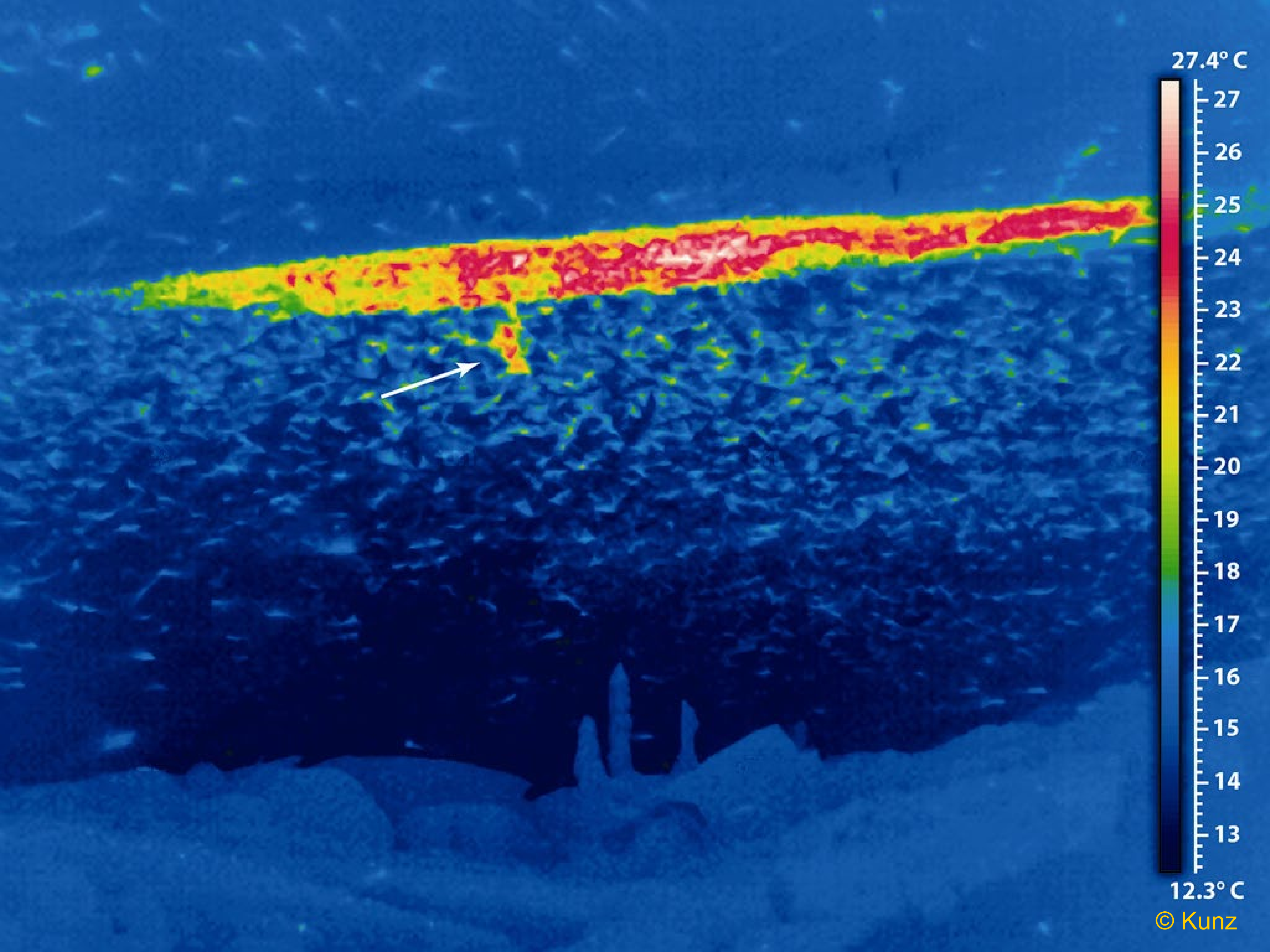

© Fuller

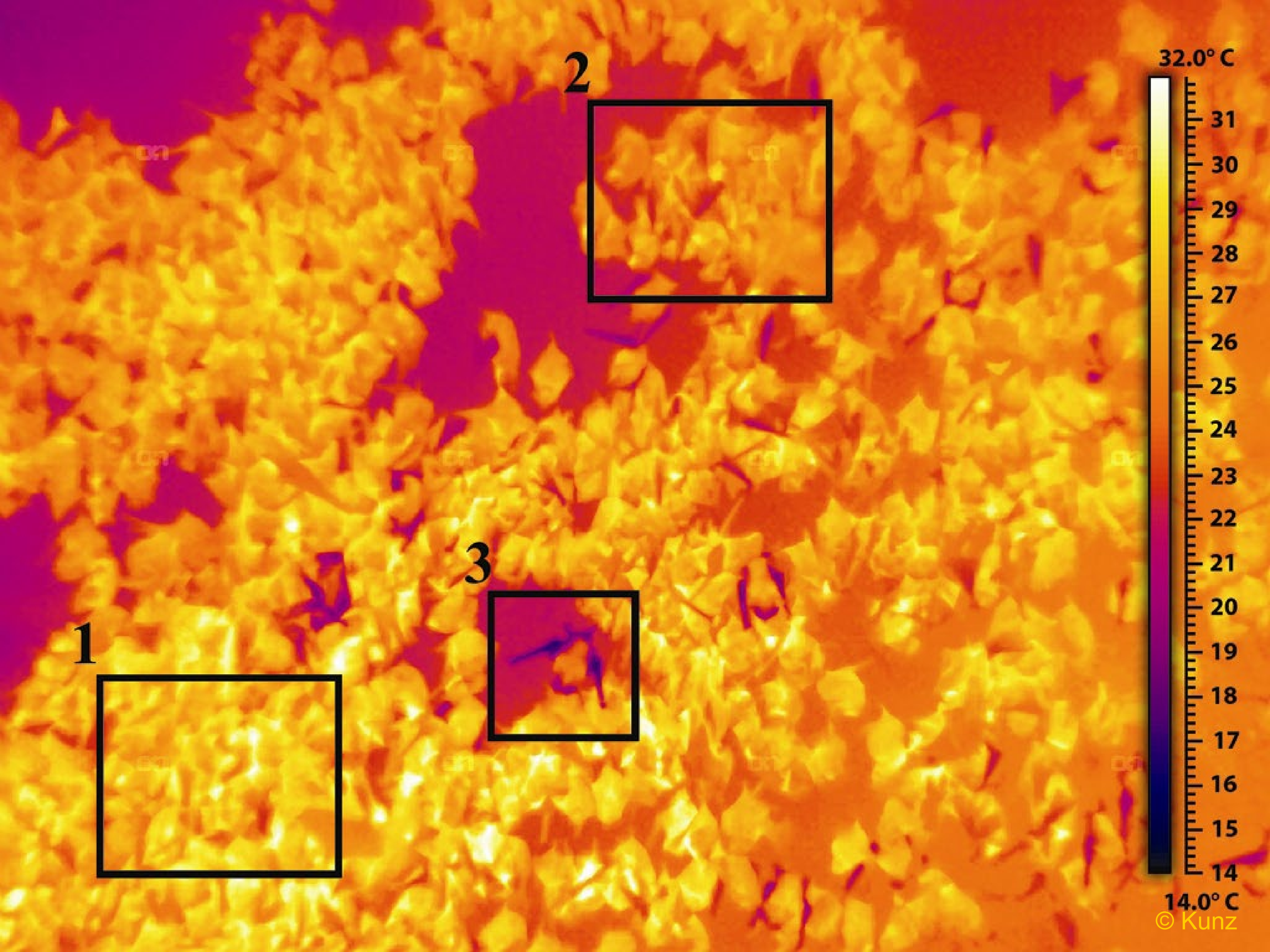


0
© Fuller

Thermal Image of Bats Roosting in Carlsbad Caverns



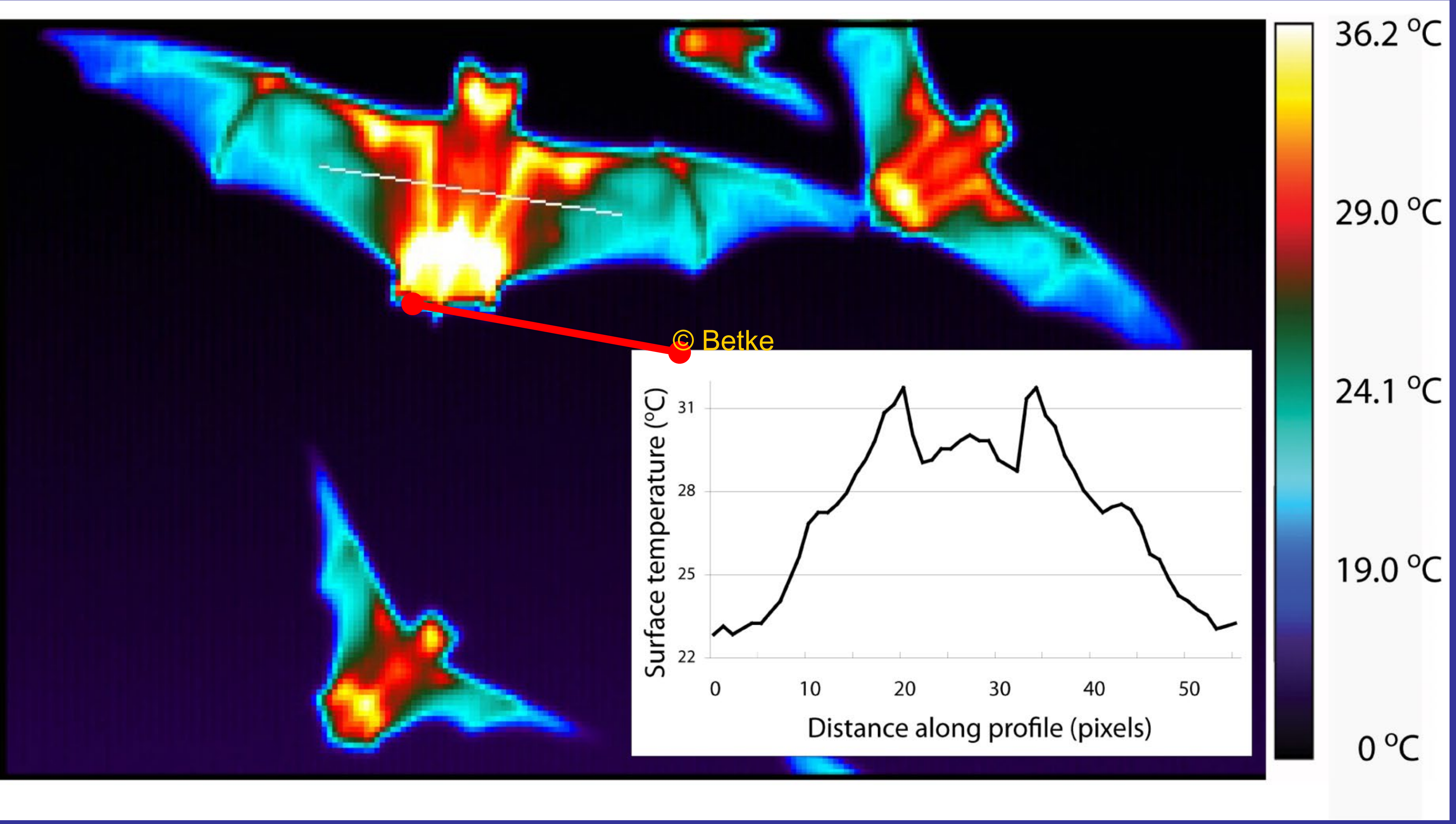

27.4° C
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12.3° C
© Kunz



32.0° C
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
14.0° C
1
2
3
© Kunz

Thermal Infrared Imaging

Physiology

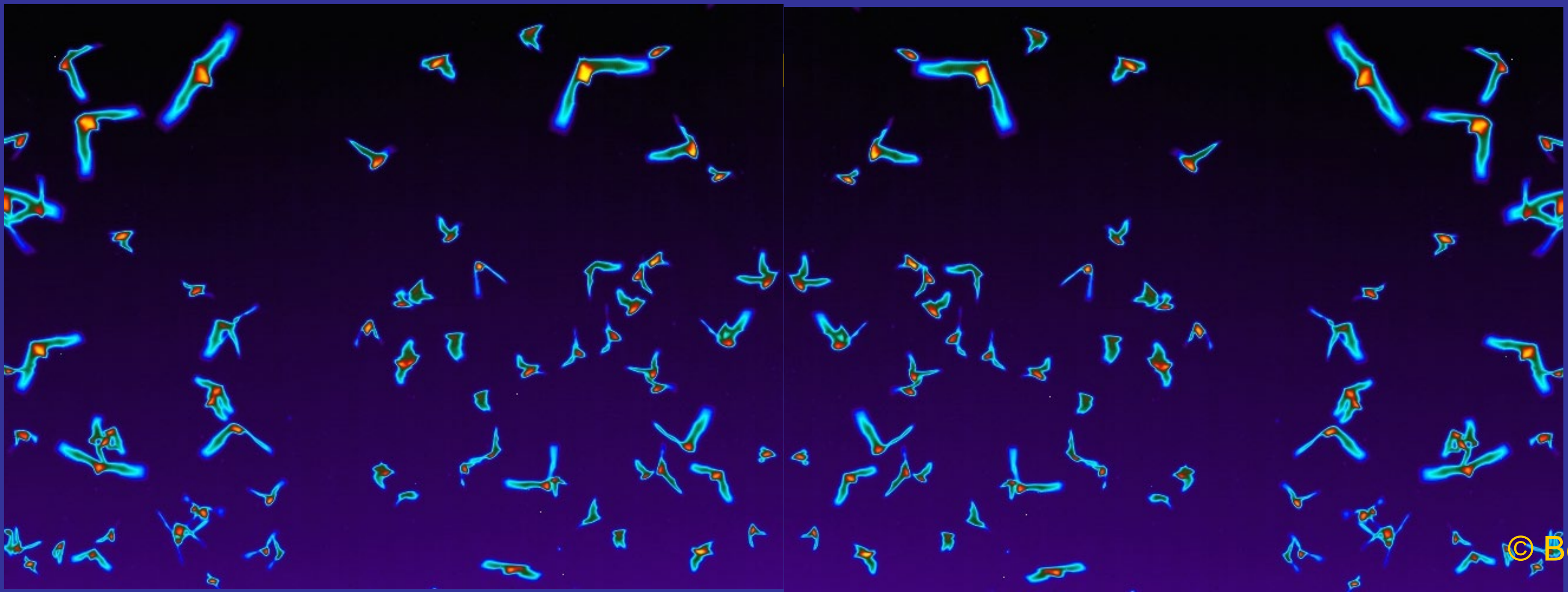


36.2 °C
29.0 °C
24.1 °C
19.0 °C
0 °C
© Betke
Surface temperature (°C)
31
28
25
22
0
10
20
30
40
50
Distance along profile (pixels)

Counting bats

- Deploy our cameras perpendicular to the flow of bats

- Record the entire emergence (~ 45 minutes)



Challenges

- Initiation of tracking automatically, anywhere in frame
- Background / foreground classification in noise and clutter
- Number of objects unknown
- Objects have similar appearance
- Previously tracked objects may not be observed in some frames due to occlusion or low signal-to-noise ratio; tracking resume as soon as objects reappear
- Scalable solution that works reliably for > 100,000 objects

Pest Control Services

- Bats eat some of the most destructive agricultural pests in the world

Corn Earworm

John L. Capinera, University of Florida

Corn Earworm Moth

http://en.wikipedia.org/wiki/Corn_earworm

- One lactating Brazilian free-tailed bat can consume up to 114 corn earworm moths in one night
- One bat saves farmers $0.02 per night in mid-June

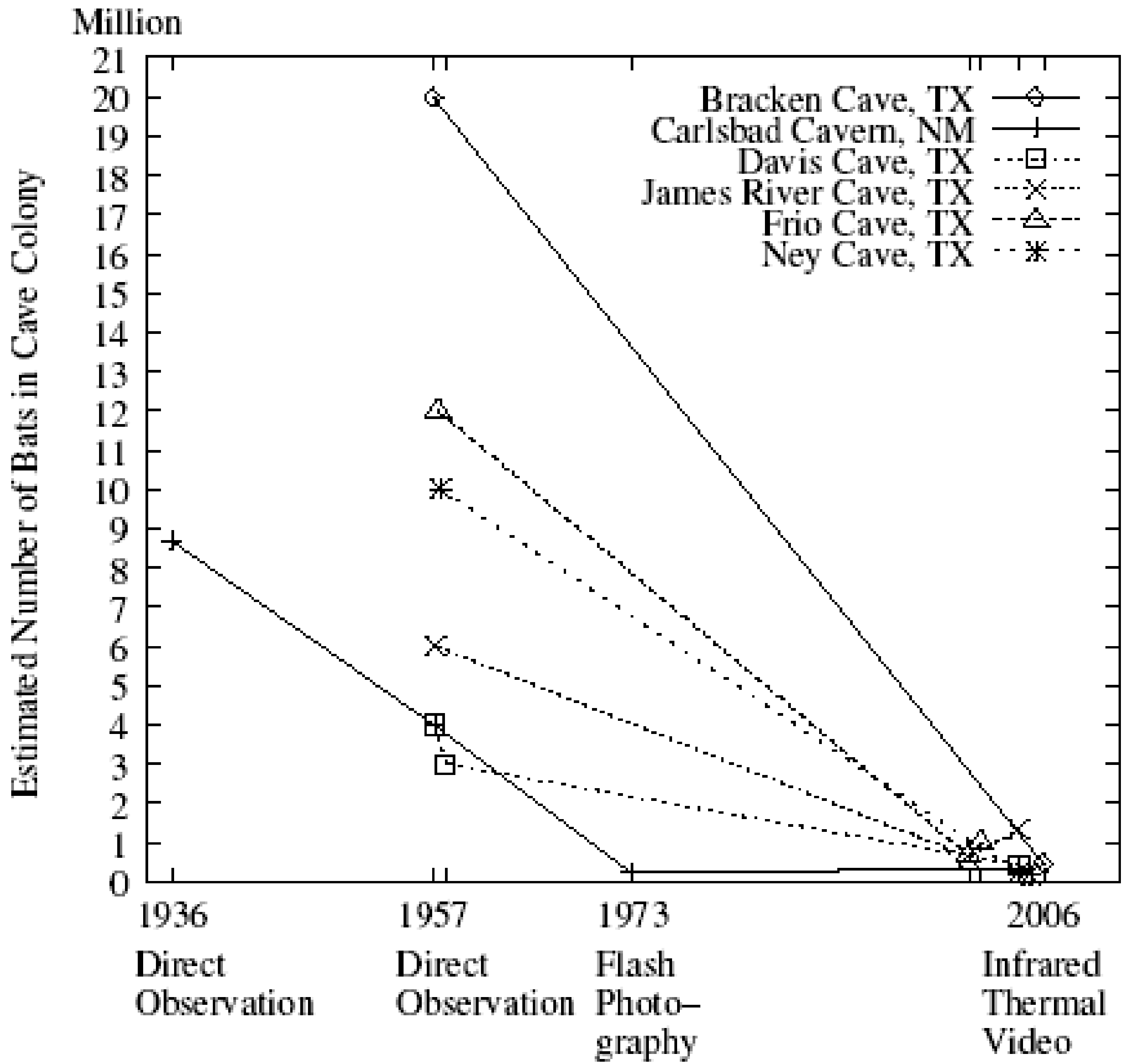
Motivation: Ecology

- How many bats do we have in North America?

  “Guesstimate” for *Tadarida brasiliensis* for 1950s:
  150 million

  Our estimate for current population:
  11 million

- Need non-invasive and accurate census methods

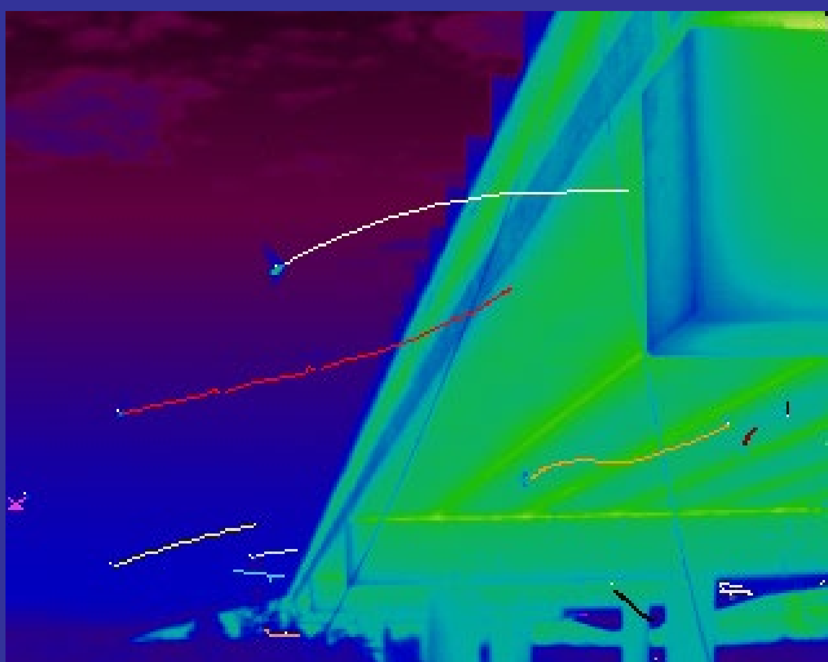


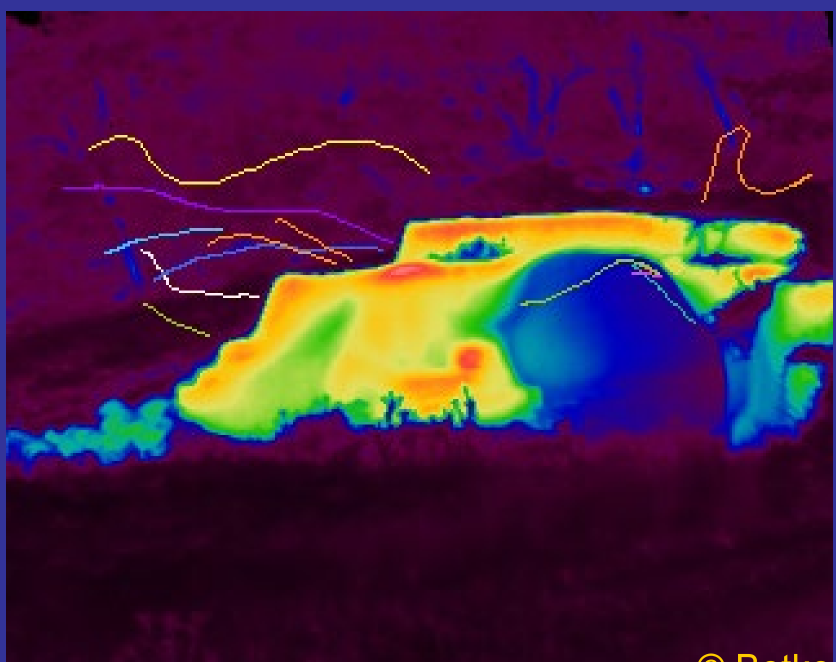
Million
Estimated Number of Bats in Cave Colony
Bracken Cave, TX
Carlsbad Cavern, NM
Davis Cave, TX
James River Cave, TX
Frio Cave, TX
Ney Cave, TX
1936
Direct Observation
1957
Direct Observation
1973
Flash Photo–graphy
2006
Infrared Thermal Video

Cross-disciplinary Impact: Ecology

Has the *Tadarida brasiliensis* population in Texas and New Mexico been in decline?

- Inter- and intra-seasonal censusing, also at bridges
- Censusing of colonies for which there are no published estimates (e.g., Selah Chiroptorium)



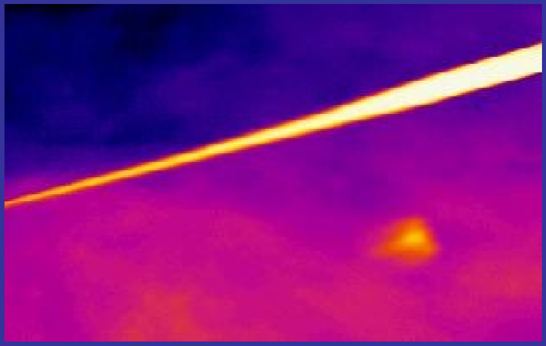


Cross-disciplinary Impact: Ecology

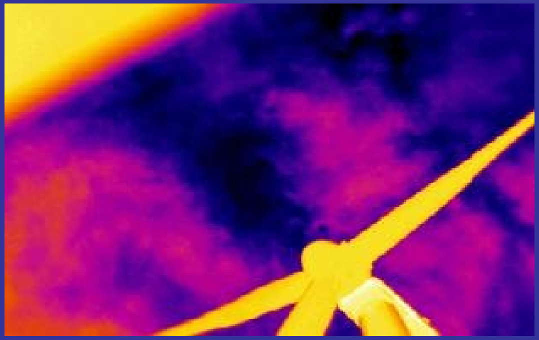
- Has the *Tadarida brasiliensis* population in Texas and New Mexico been in decline?

  Our research showed decline.

- Impact of wind energy parks



- Animal behavior and interaction:
  - Need video analysis methods for studying foraging habits
  - Need analysis methods to study flight behavior



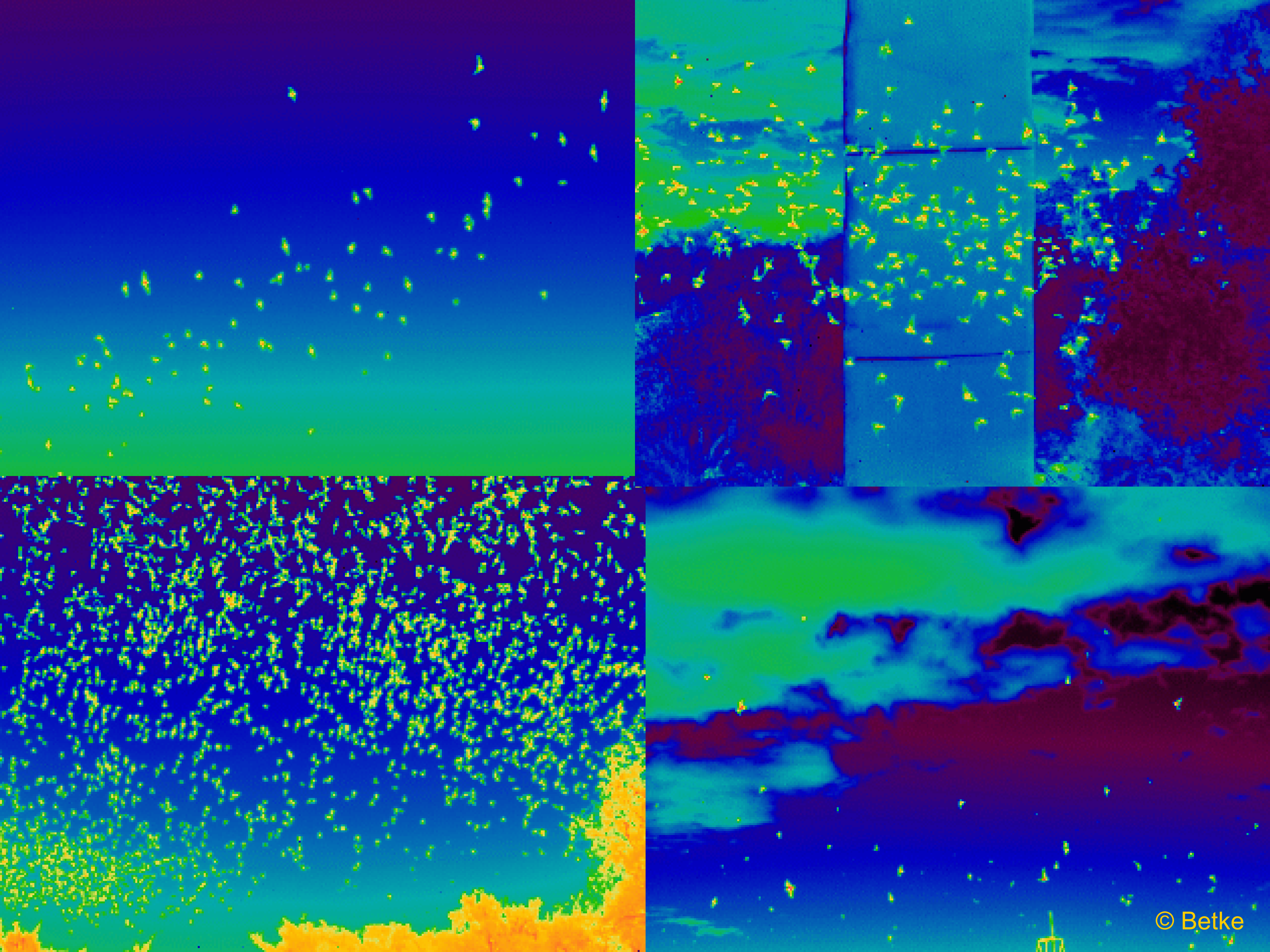
Publications

http://www.cs.bu.edu/fac/betke

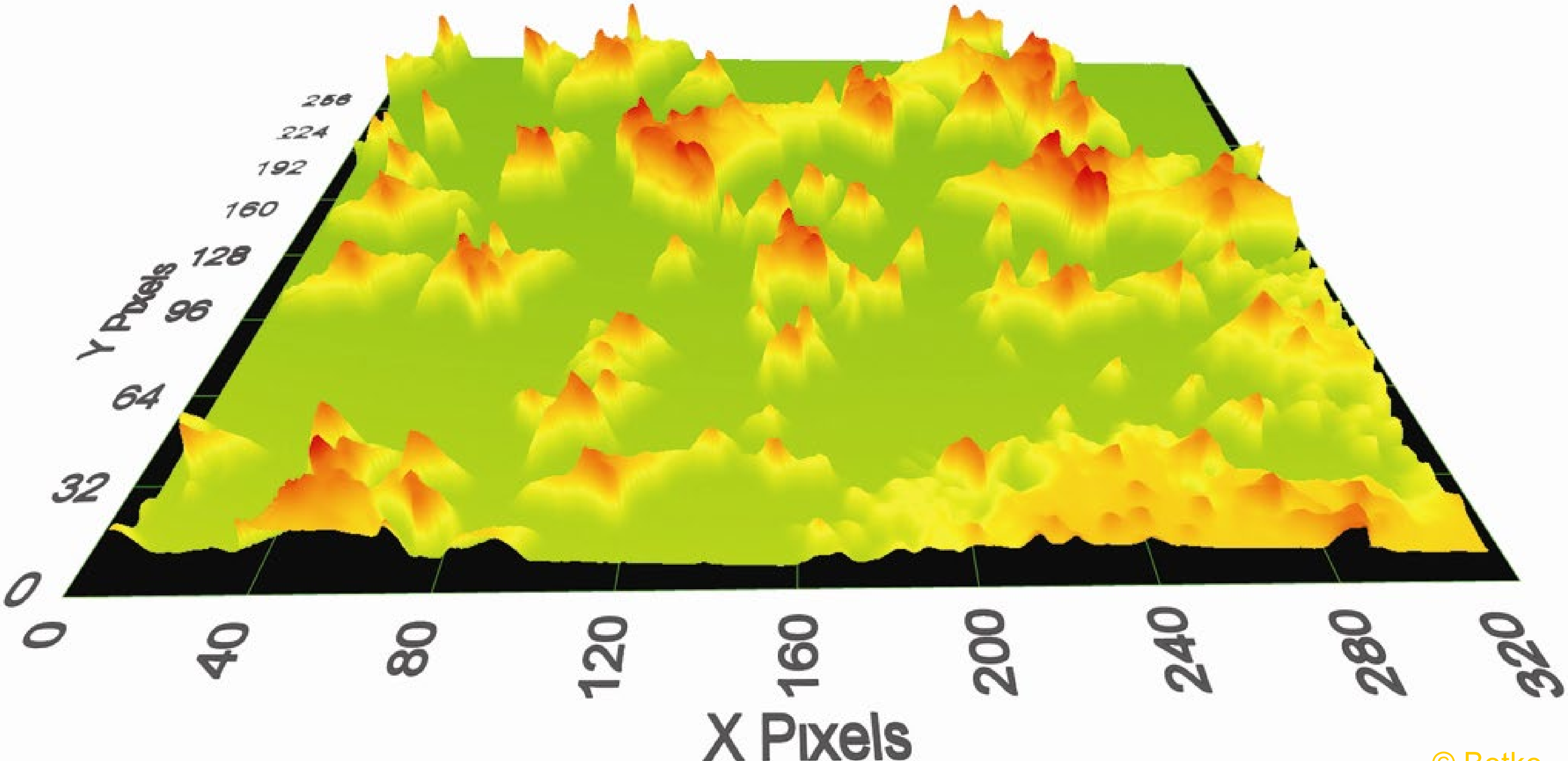
- IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)
- International Conference on Computer Vision (ICCV)
- PhD & MA Theses in Computer Science at Boston University
- Annual North American Symposium on Bat Research
- Frontiers in Ecology
- Journal of Mammalogy
- Nature, Research Highlights & Nature Scientific Reports

- Our own book on multi-object, multi-view tracking: 2016

SAMSUNG


© Betke

Thermal Intensities of Emerging Bats



Detection Method

1. Adaptive Pixel-based Filter

   Dynamic Gaussian models of brightness changes.

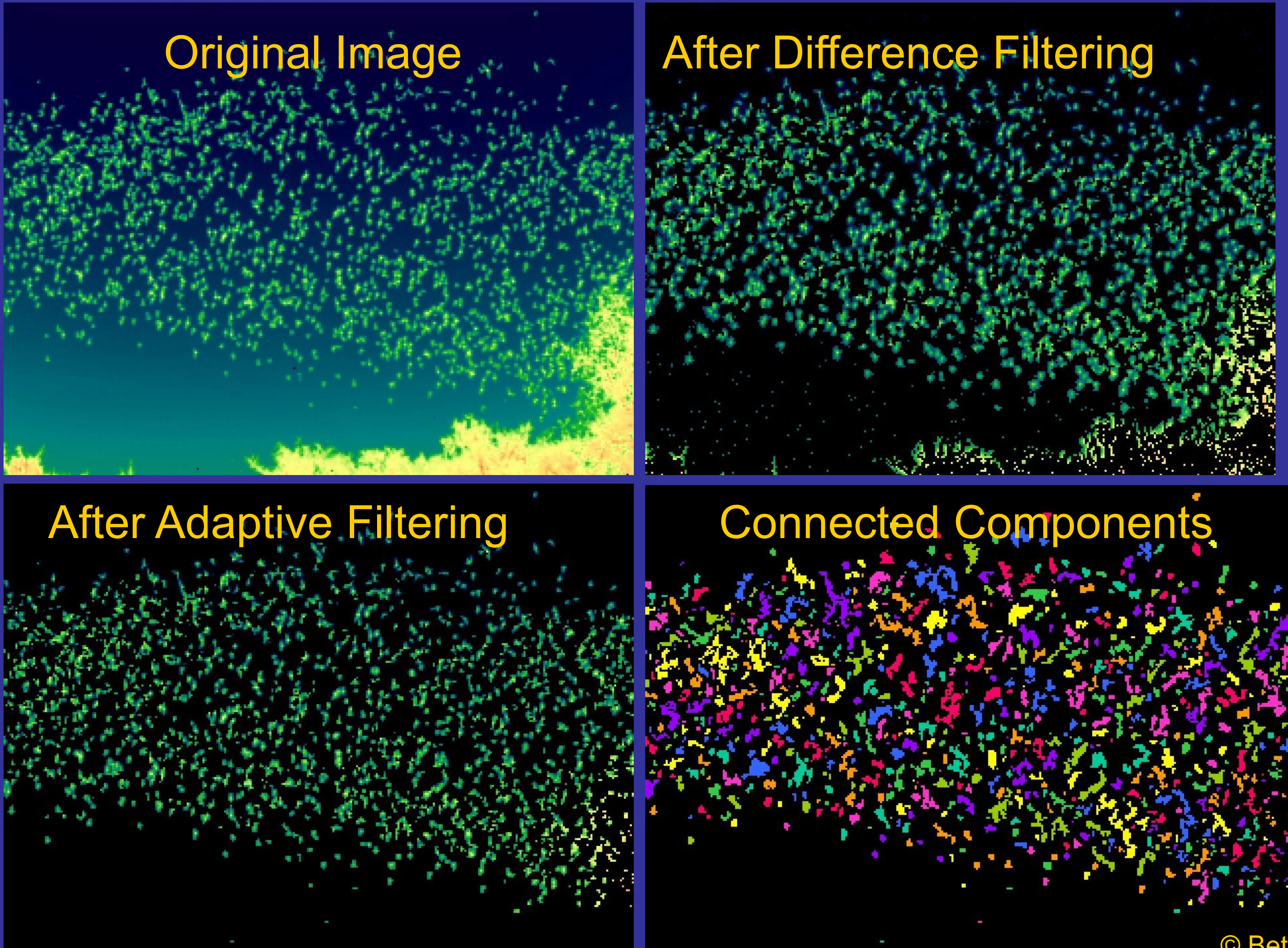
   Filter

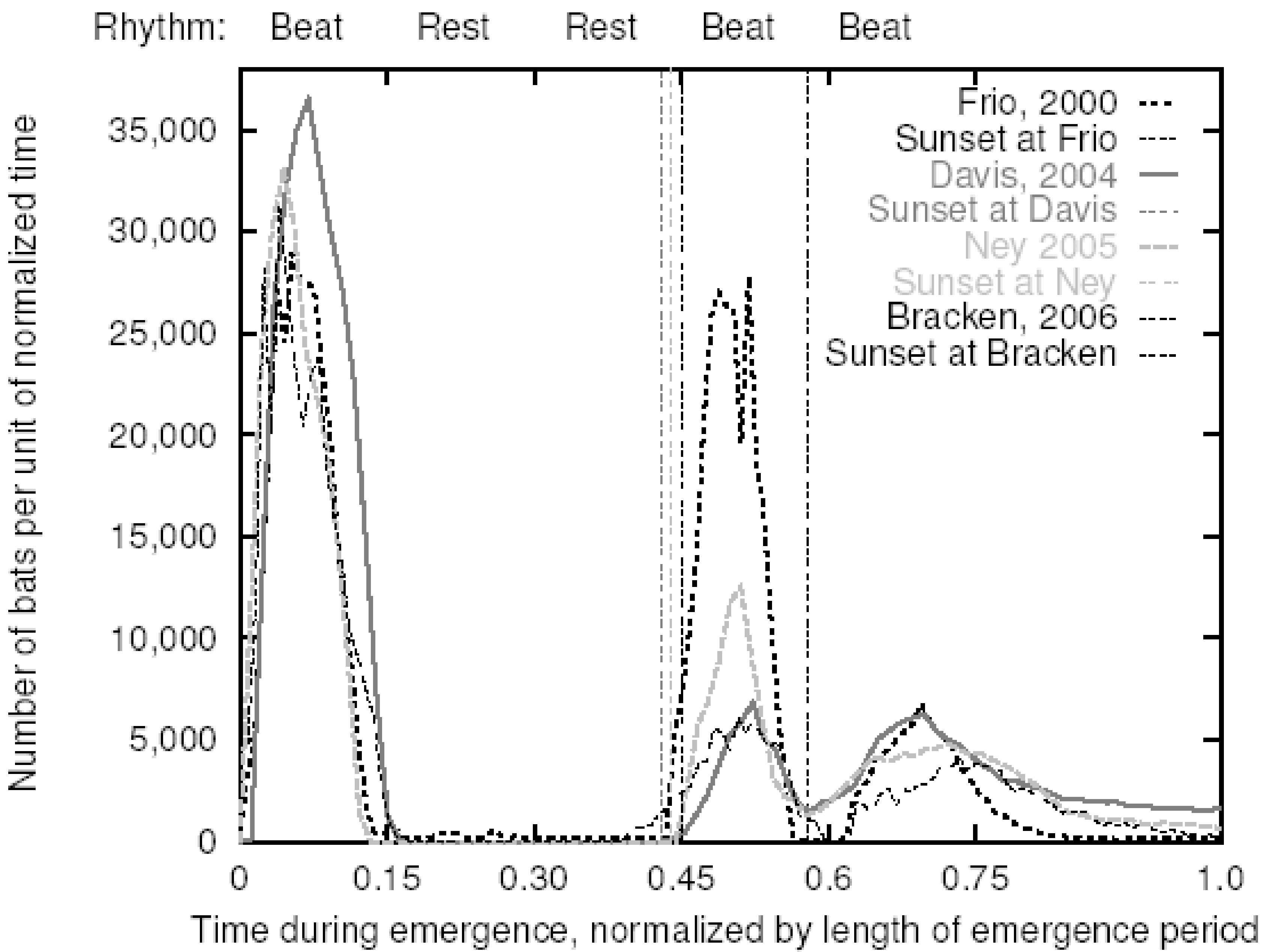
   if $k\,\sigma(x,y,t) < |\,I(x,y,t) - \mu(x,y,t)|$

   for k = 5%; mean μ and std. dev. σ updated in time window

2. Region Analysis:

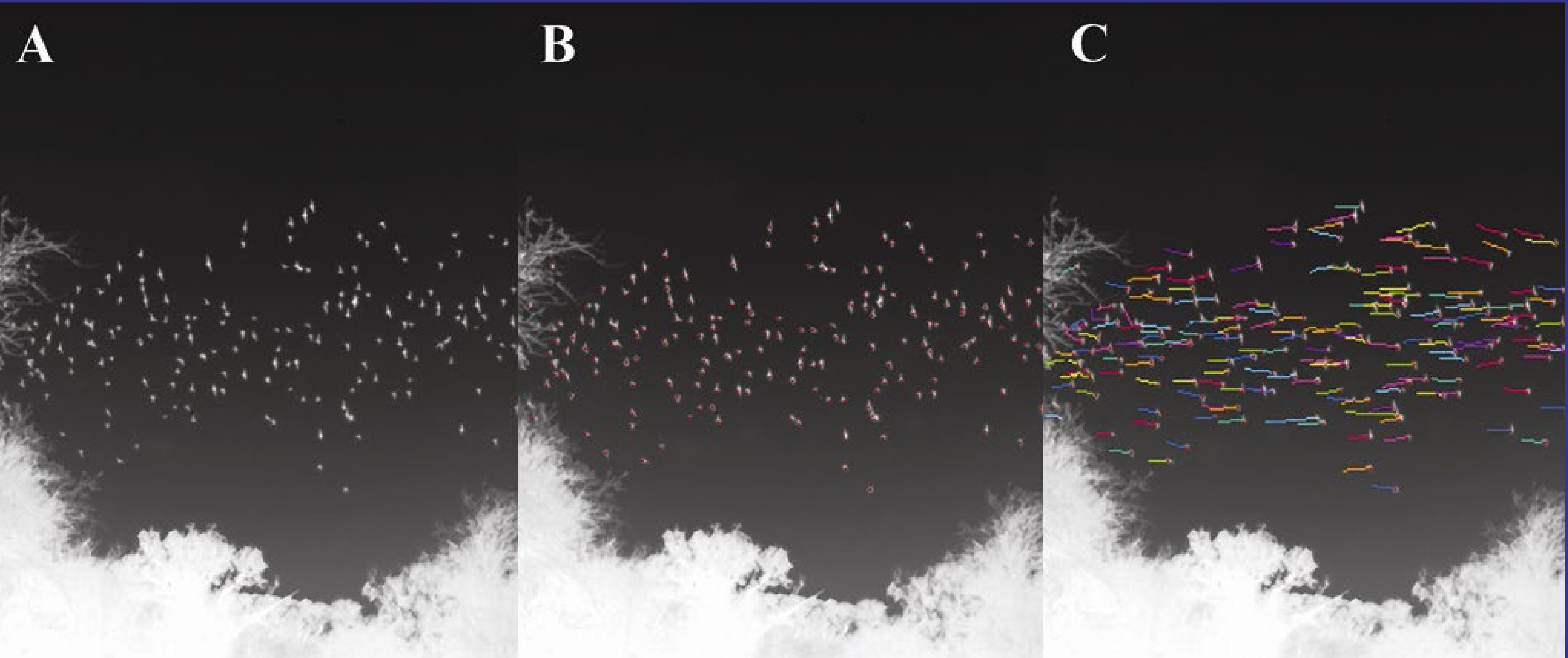
   Intensity peaks within connected components of processed pixels


Original Image
After Difference Filtering
After Adaptive Filtering
Connected Components

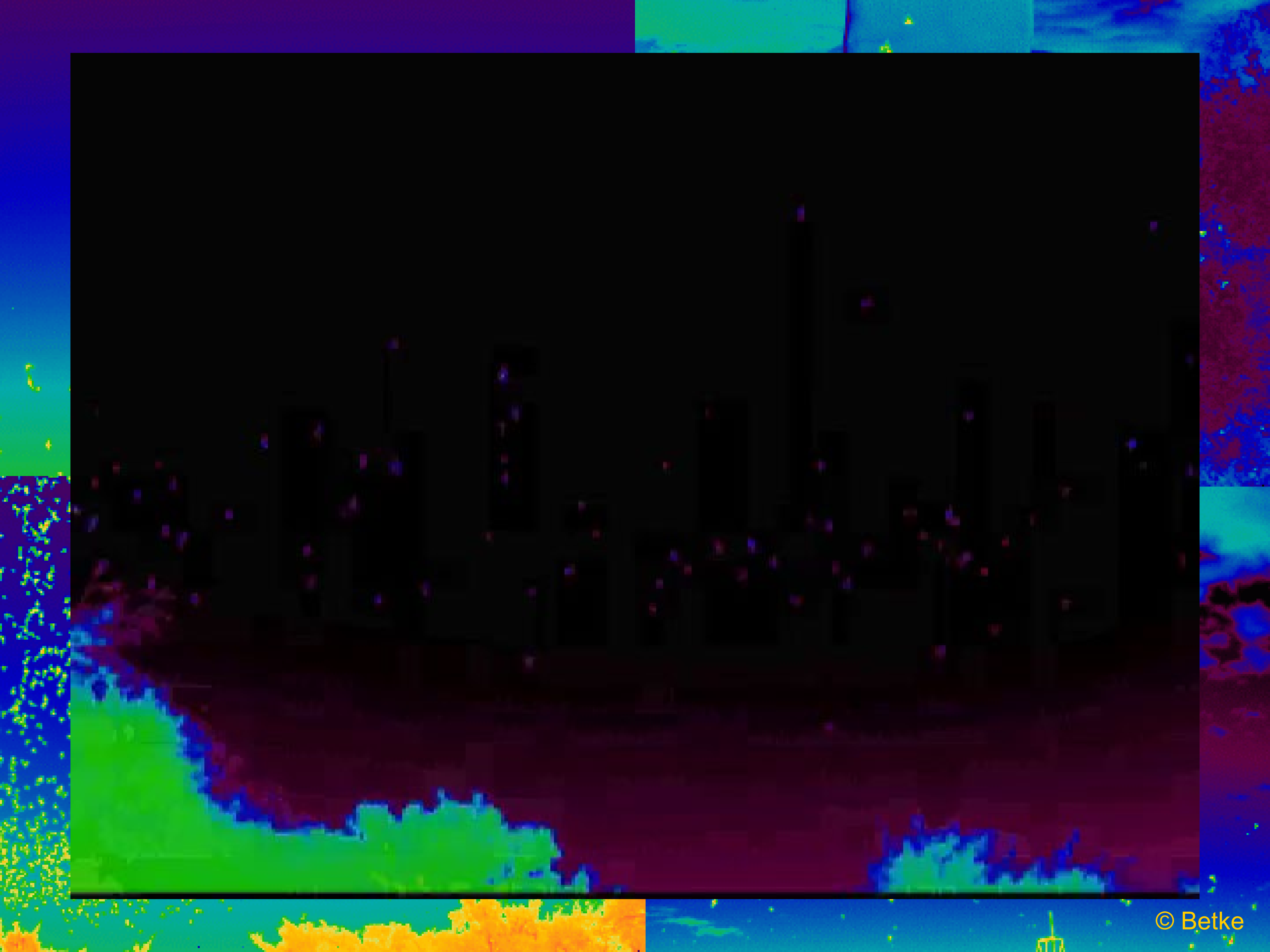

Rhythm: Beat Rest Rest Beat Beat
Frio, 2000
Sunset at Frio
Davis, 2004
Sunset at Davis
Ney 2005
Sunset at Ney
Bracken, 2006
Sunset at Bracken
Number of bats per unit of normalized time
35,000
30,000
25,000
20,000
15,000
10,000
5,000
0
0
0.15
0.30
0.45
0.6
0.75
1.0
Time during emergence, normalized by length of emergence period

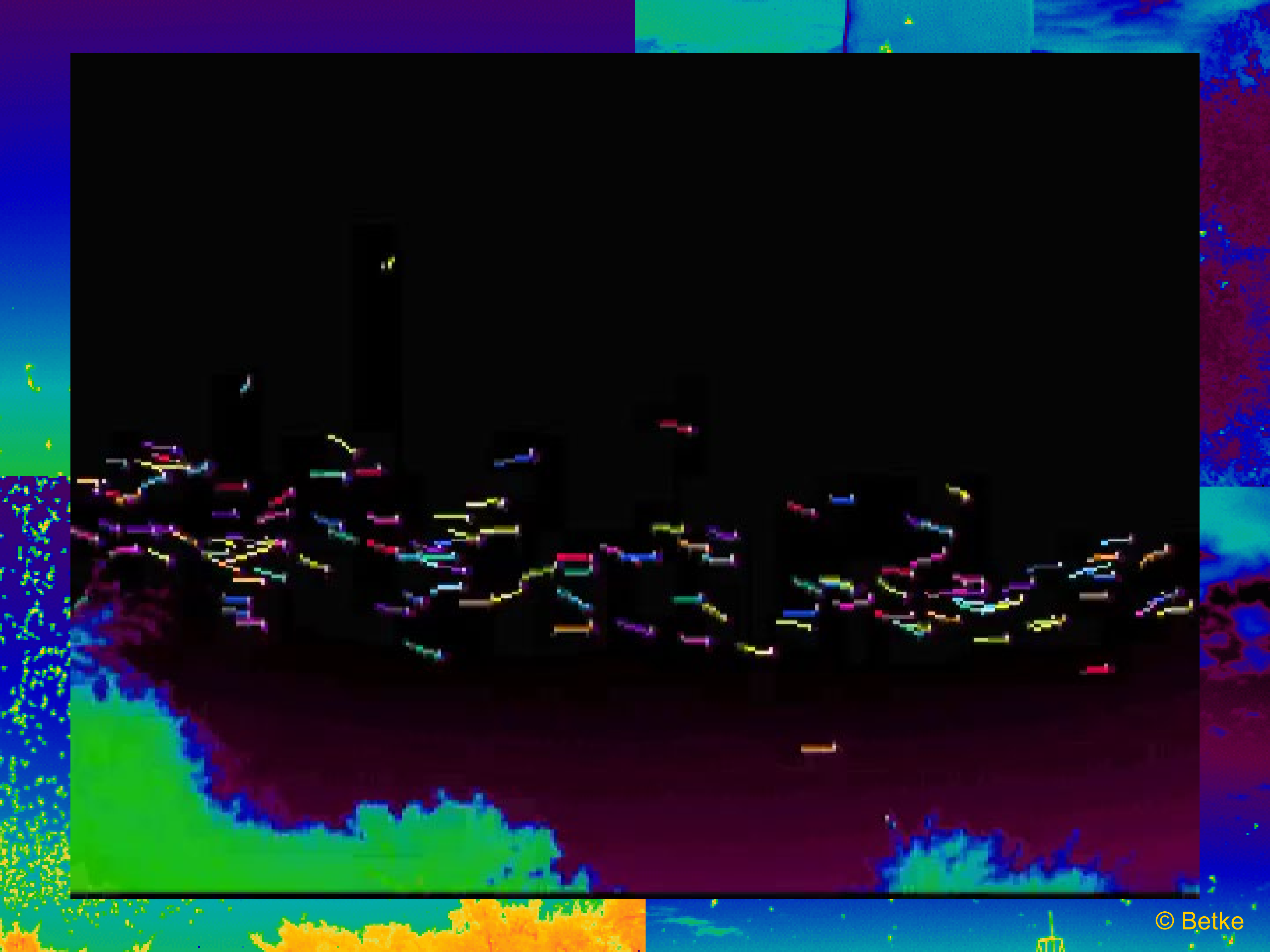
So far: Counting bats

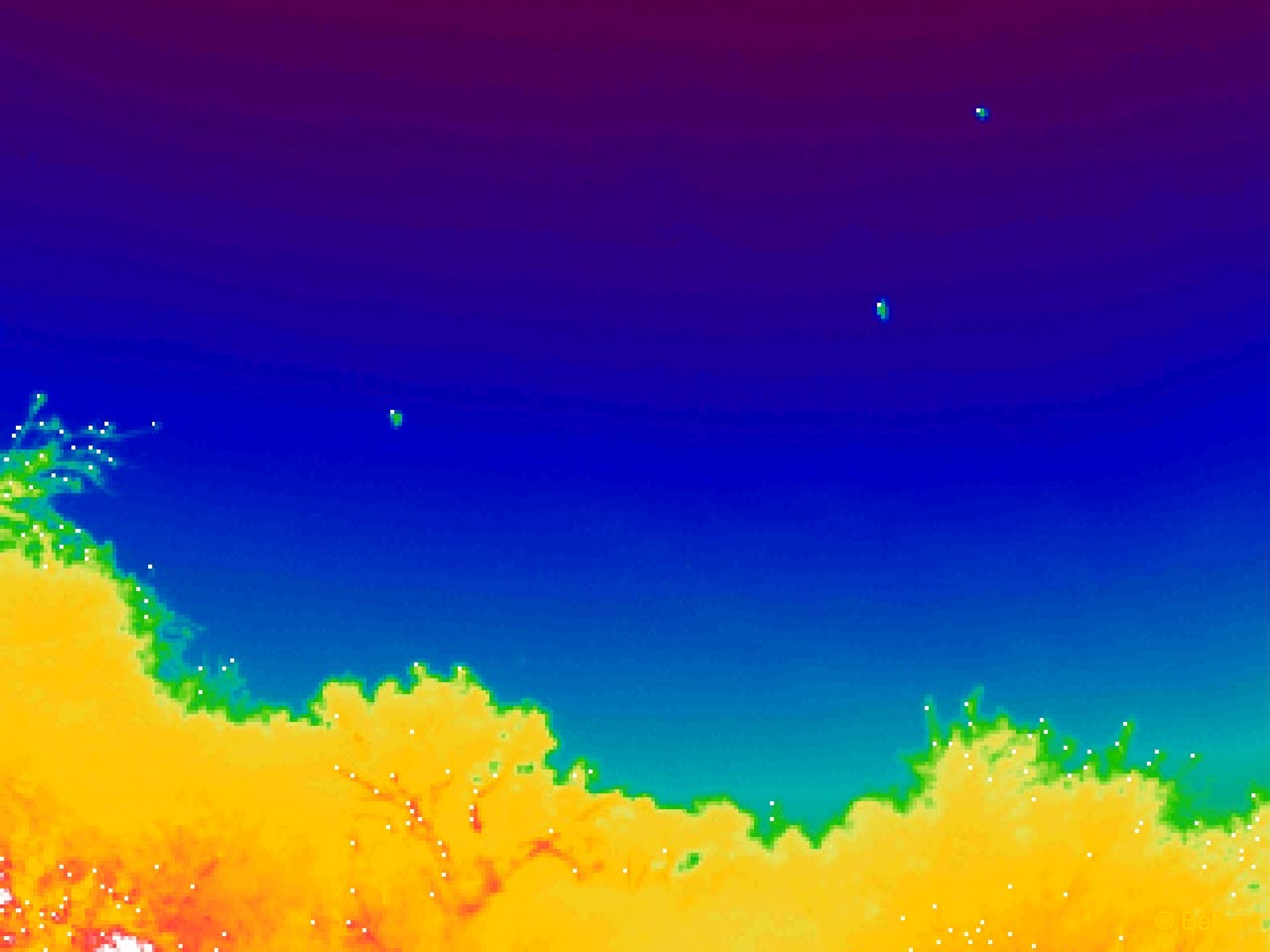
Now: Tracking bats



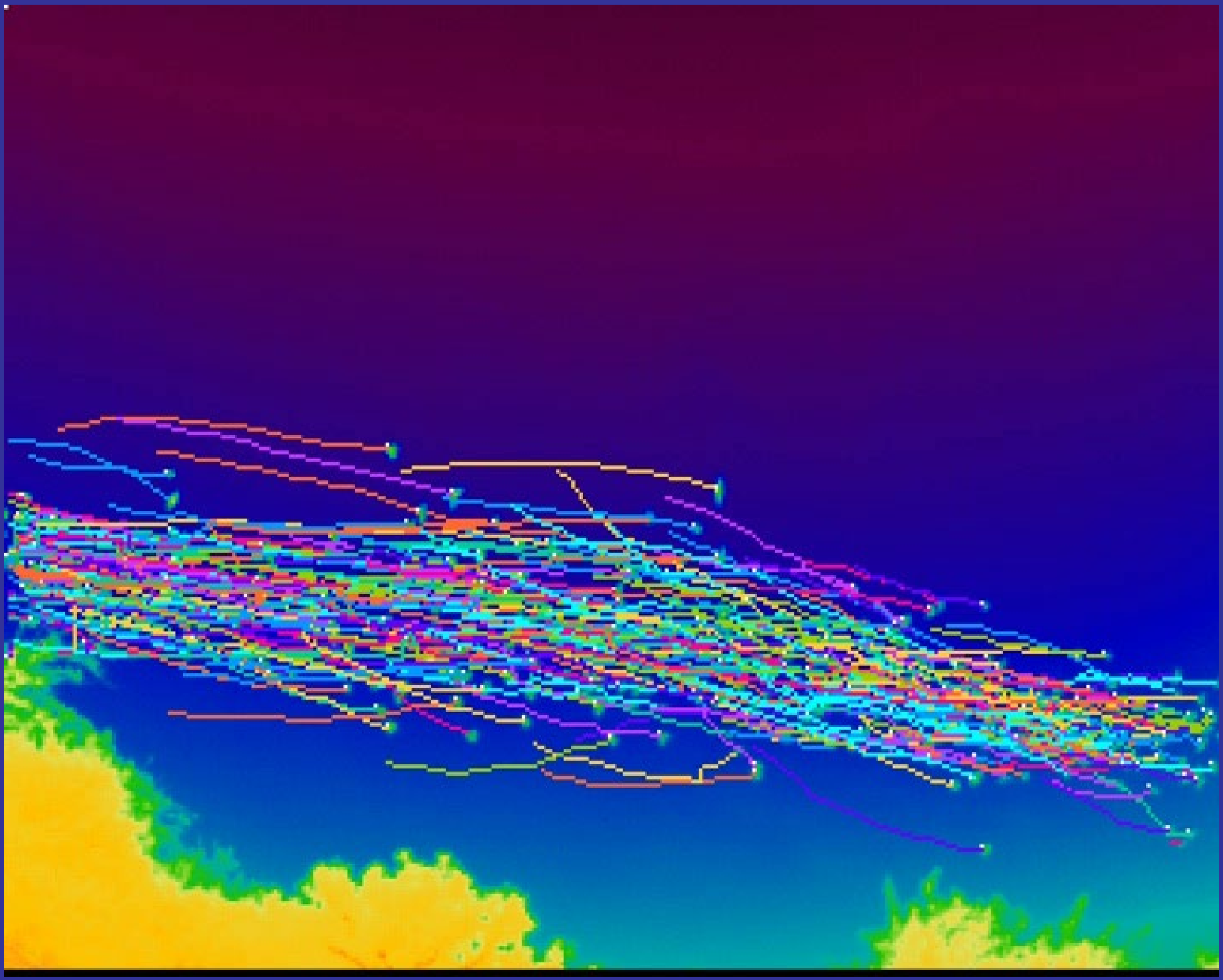
A. Original thermal image (2 bytes / pixel, not shown with false color)
B. Detected bats via brightness peak (red)
C. Tracked bats





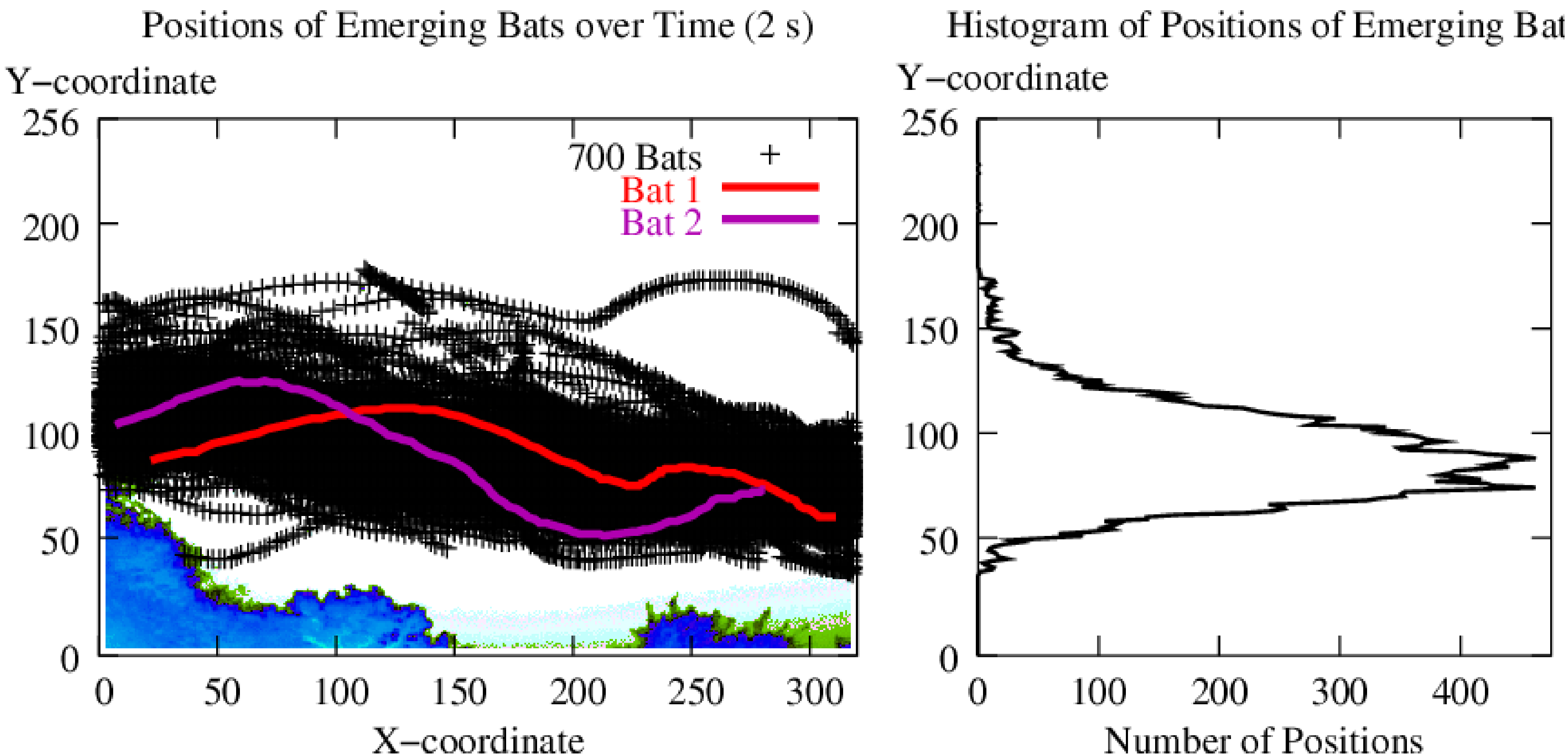

© Betke

Tracking Bats Emerging in a Column Formation from a Cave in Texas

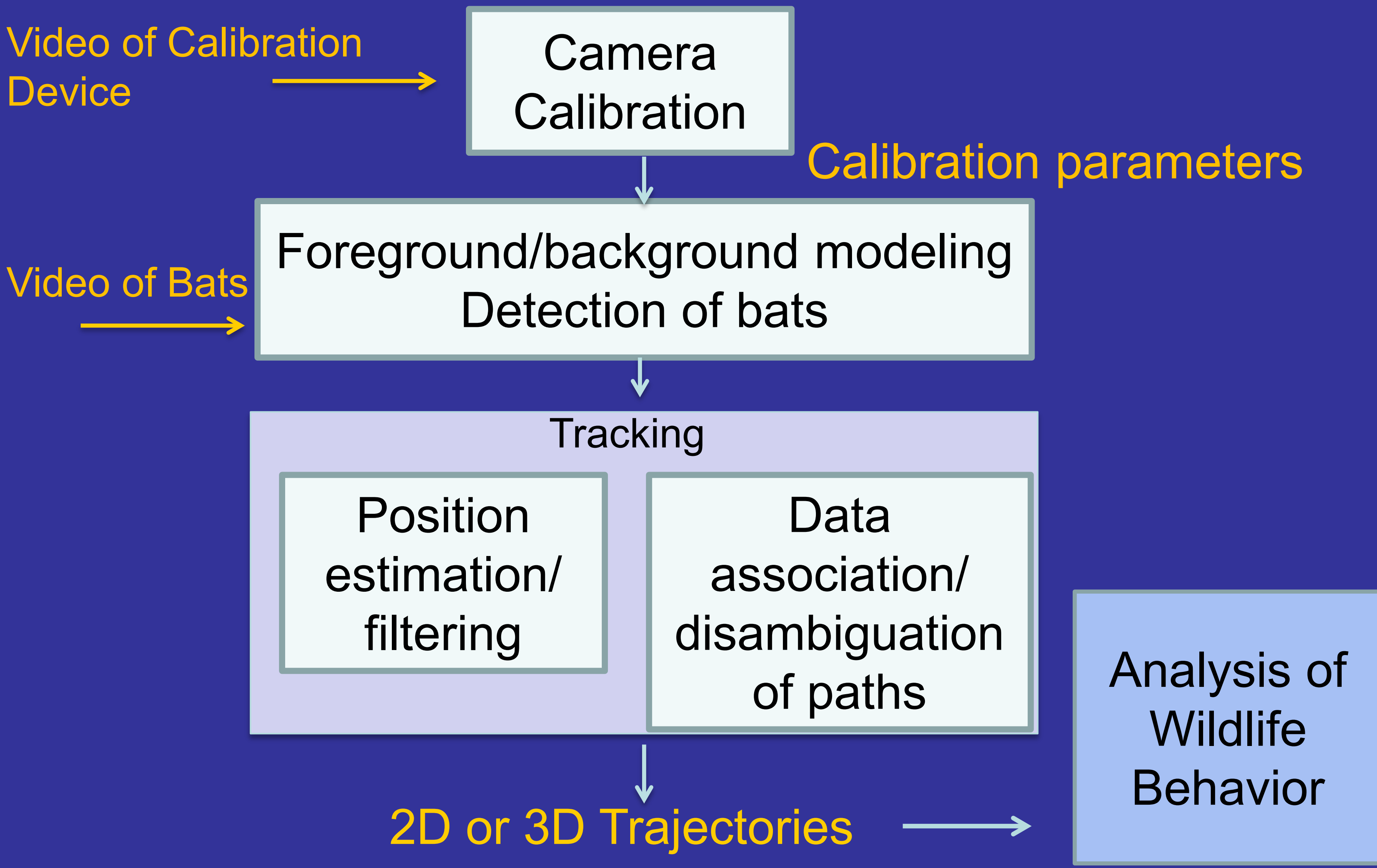


Opportunities for Studying Wildlife

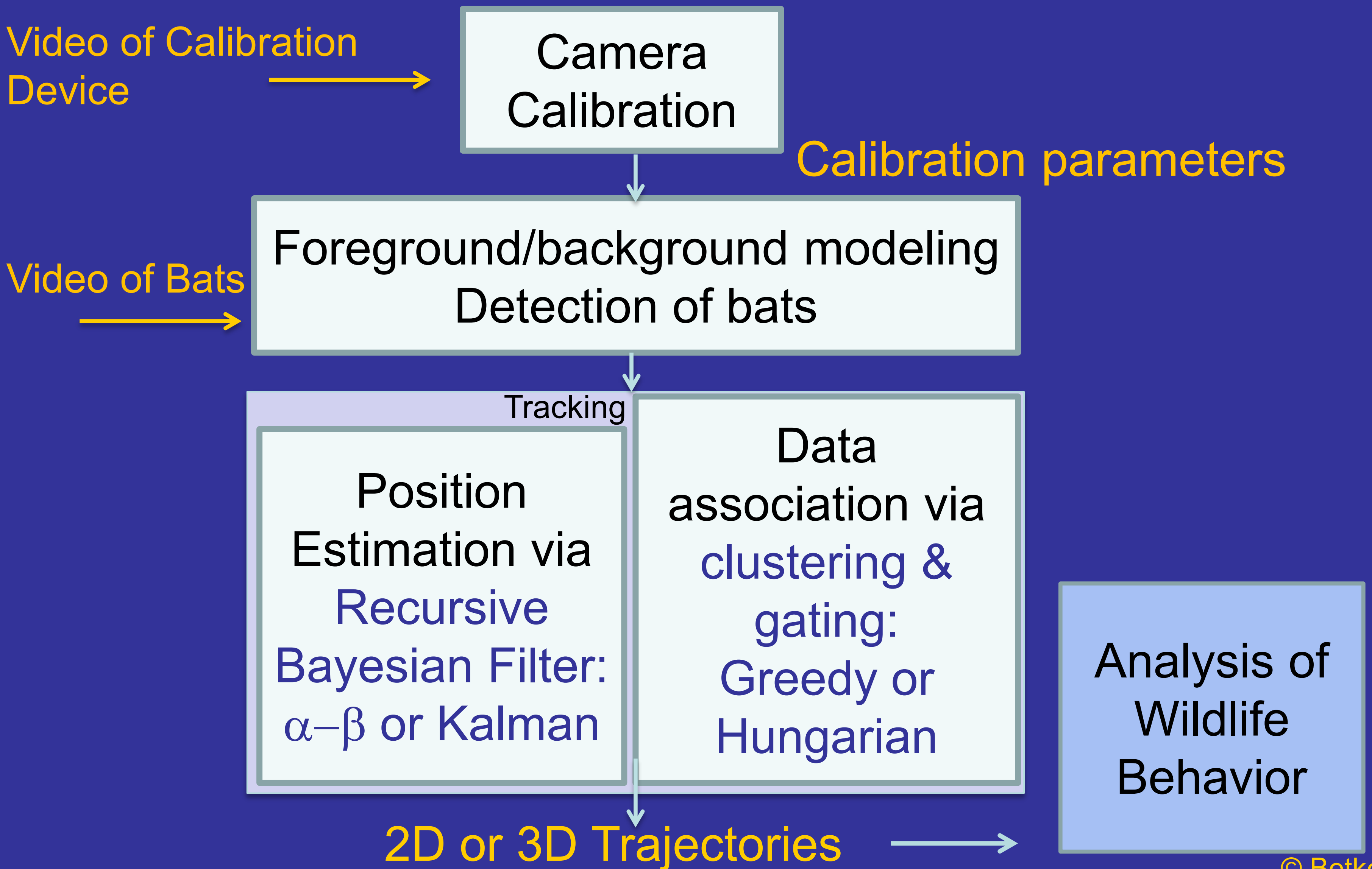
How do bats fly in groups?



Computer Vision Workflow



Computer Vision Workflow



Video-based Tracking

Measurement of object in image at time *t*: *z(t)*

e.g. *z(t)* is observed property of object in image, e.g.:

- horizontal & vertical position in image
- brightness (grey level)
- circularity

State of object at time *t*: *x(t), e.g., 3D position in world*

Video-based Tracking

Measurement of object in image at time *t*: *z(t)*

e.g. *z(t)* is observed property of object in image, e.g.:

- horizontal & vertical position in image
- brightness (grey level)
- circularity

z(t) depends on x(t)

State of object at time *t*: *x(t), e.g., 3D position in world*

x(t) depends on x(t-1)

Video-based Tracking: Notation

Measurement of object in image at time *t*: *z(t)*
z(t) depends on x(t)

State of object at time *t*: *x(t), e.g., 3D position in world*
x(t) depends on x(t-1)

$\hat{x}(t|t)$ = *current state estimate based on t measurements*
$\hat{x}(t|t-1)$ = *current state estimate based on t-1 measurements*
$\hat{x}(t-1|t-1)$ = *previous state estimate based on all the measurements that were available at the time*

Video-based Tracking: Notation

Measurement of object in image at time t: $z(t)$
$z(t)$ depends on $x(t)$

State of object at time t: $x(t)$, *e.g., 3D position in world*
$x(t)$ depends on $x(t-1)$

$\hat{x}(t|t)$ = *current state estimate based on t measurements*
$\hat{x}(t|t-1)$ = *current state estimate based on t-1 measurements*
$\hat{x}(t-1|t-1)$ = *previous state estimate based on all the measurements that were available at the time*

Welch & Bishop's notation, link to Kalman filter intro on course website:
$\hat{x}(t|t) = \hat{x}_t$
$\hat{x}(t|t-1) = \hat{x}_t^-$
$\hat{x}(t-1|t-1) = \hat{x}_{t-1}^-$

Wikipedia notation: $\hat{x}_{t|t}$, $\hat{x}_{t|t-1}$, $\hat{x}_{t-1|t-1}$

Video-based Tracking

Measurement of object in image at time *t*: $z(t)$

e.g. $z(t)$ is observed property of object in image, e.g.:

- horizontal & vertical position in image
- brightness (grey level)
- circularity

1D Measurement equation : $z(t) = x(t) + w(t)$

State of object at time *t*: $x(t)$, noise process $w(t)$

e.g., $x(t)$ can be the position, velocity, and acceleration

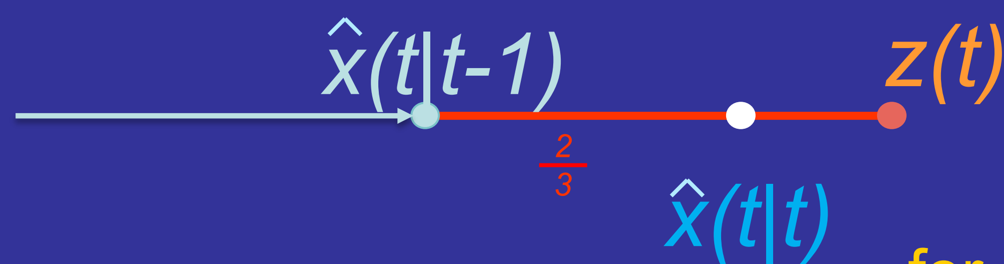
1D State equation: $x(t) = a\, x(t-1) + u(t)$

with noise process u(t) and known constant *a*

Recursive Bayesian Filter: Kalman

- Kalman filter minimizes Bayesian mean square error
  $E[\,(x(t) - \hat{x}(t|t))^2\,]$:

Estimate update: $\hat{x}(t|t) = \hat{x}(t|t-1) + K(t)\,\left(z(t) - \hat{x}(t|t-1)\right)$



for $K(t)=2/3$

Recursive Bayesian Filters

- Kalman filter minimizes Bayesian mean square error

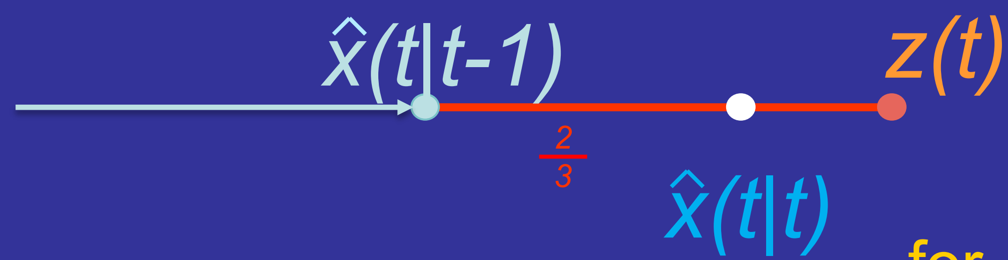
  $E[\ (x(t) - \hat{x}(t|t)\ )^2\ ]$:

  Estimate update: $\hat{x}(t|t) = \hat{x}(t|t-1) + K(t)\ \ (z(t) - \hat{x}(t|t-1))$

- Alpha-beta filter keeps track of position & velocity:

  Estimate update: $\hat{x}(t|t) = \hat{x}(t|t-1) + \alpha\ \ \ \ (z(t) - \hat{x}(t|t-1))$

  $\hat{v}(t|t) = \hat{v}(t|t-1) + \beta/\Delta T\ (z(t) - \hat{x}(t|t-1))$



for α=$K(t)$=2/3

ΔT=time between measurements.

Velocity is assumed constant

Recursive Bayesian Filters

- Kalman filter minimizes Bayesian mean square error

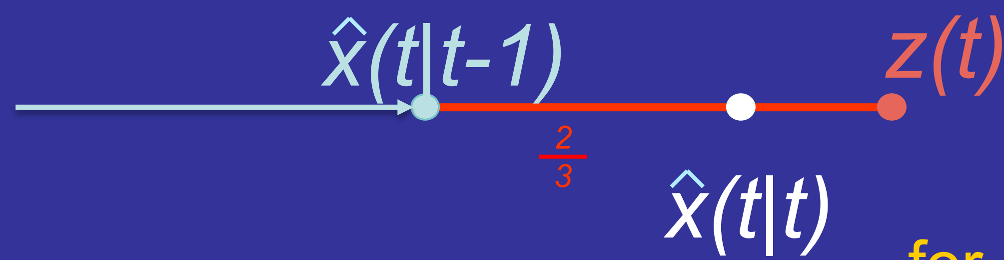
$$E[\ (x(t) - \hat{x}(t|t)\ )^2\ ]:$$

Estimate update: $\hat{x}(t|t) = \hat{x}(t|t-1) + K(t)\ \ (z(t) - \hat{x}(t|t-1))$

- Alpha-beta filter keeps track of position & velocity:

Estimate update: $\hat{x}(t|t) = \hat{x}(t|t-1) + \alpha\ \ (z(t) - \hat{x}(t|t-1))$

$$\hat{v}(t|t) = \hat{v}(t|t-1) + \beta/\Delta T\ (z(t) - \hat{x}(t|t-1))$$



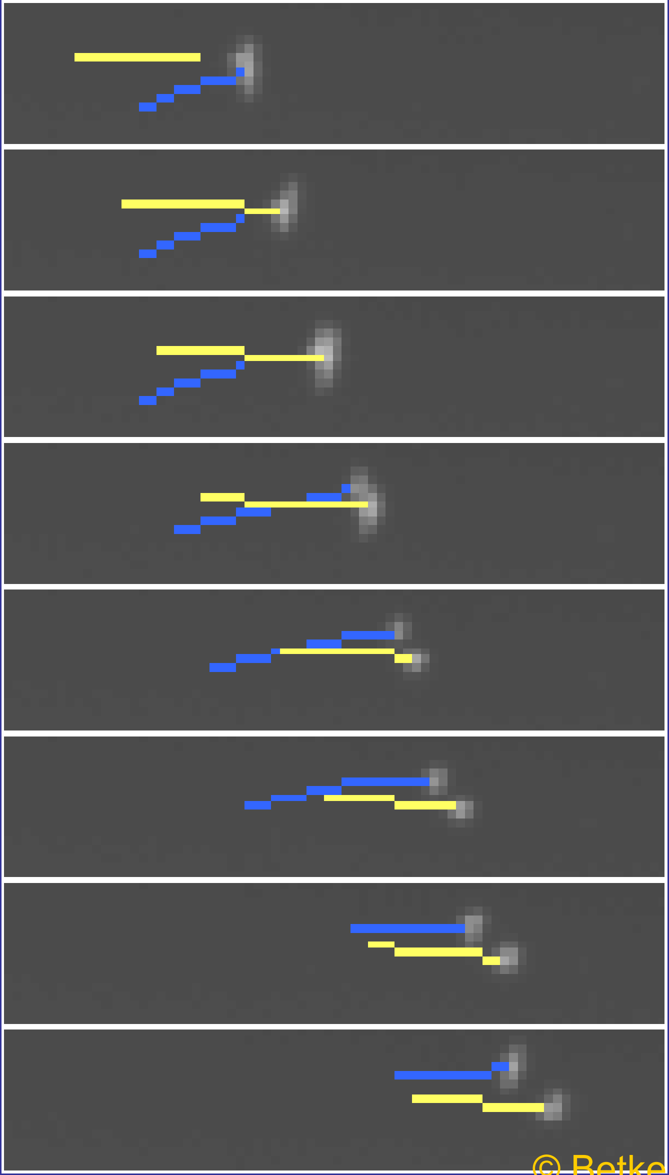
for $\alpha = K(t) = 2/3$

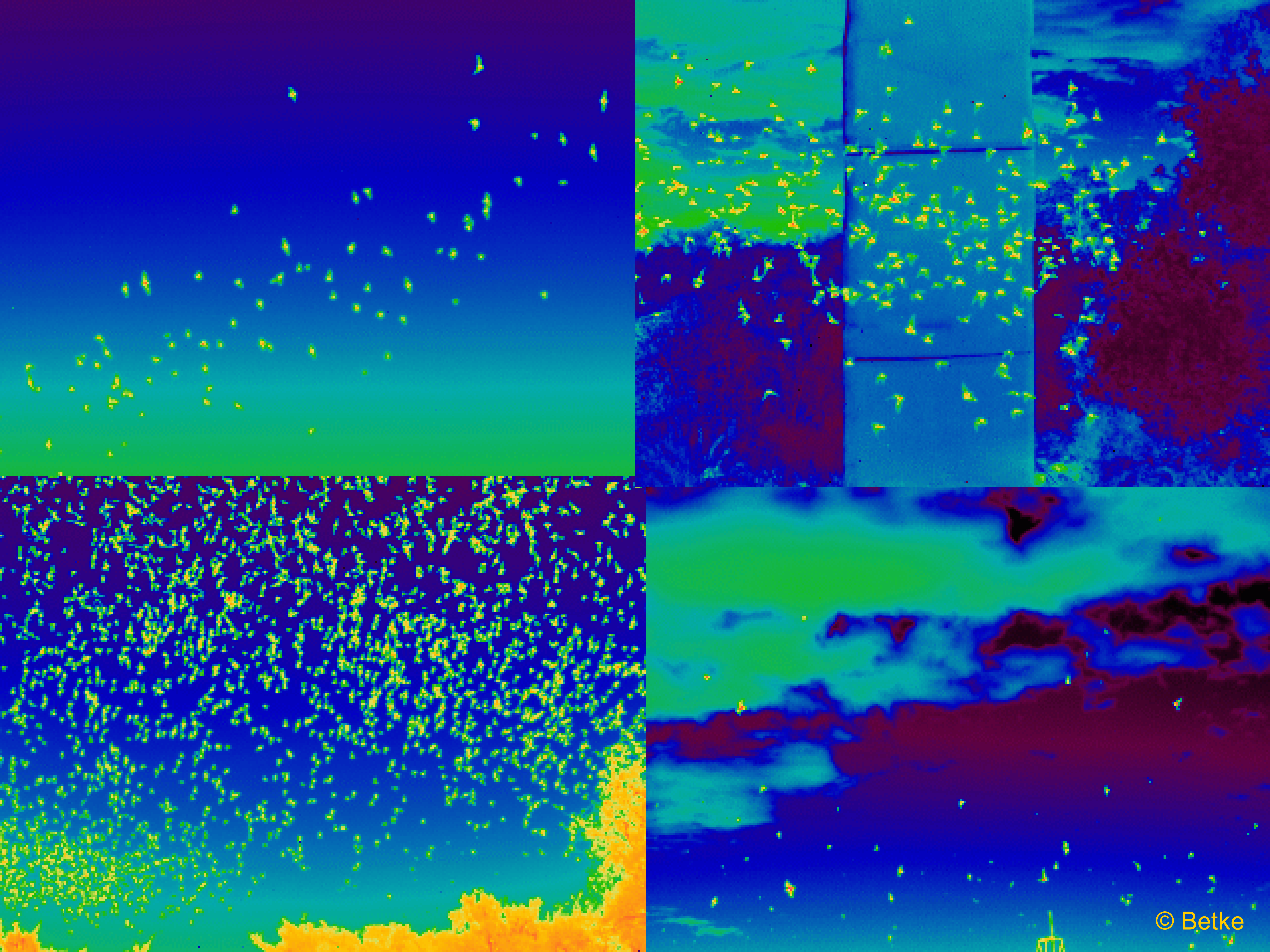
Or: *Tracking by Detection* only: states=measurements, no recursion: $\hat{x}(t|t) = z(t)$

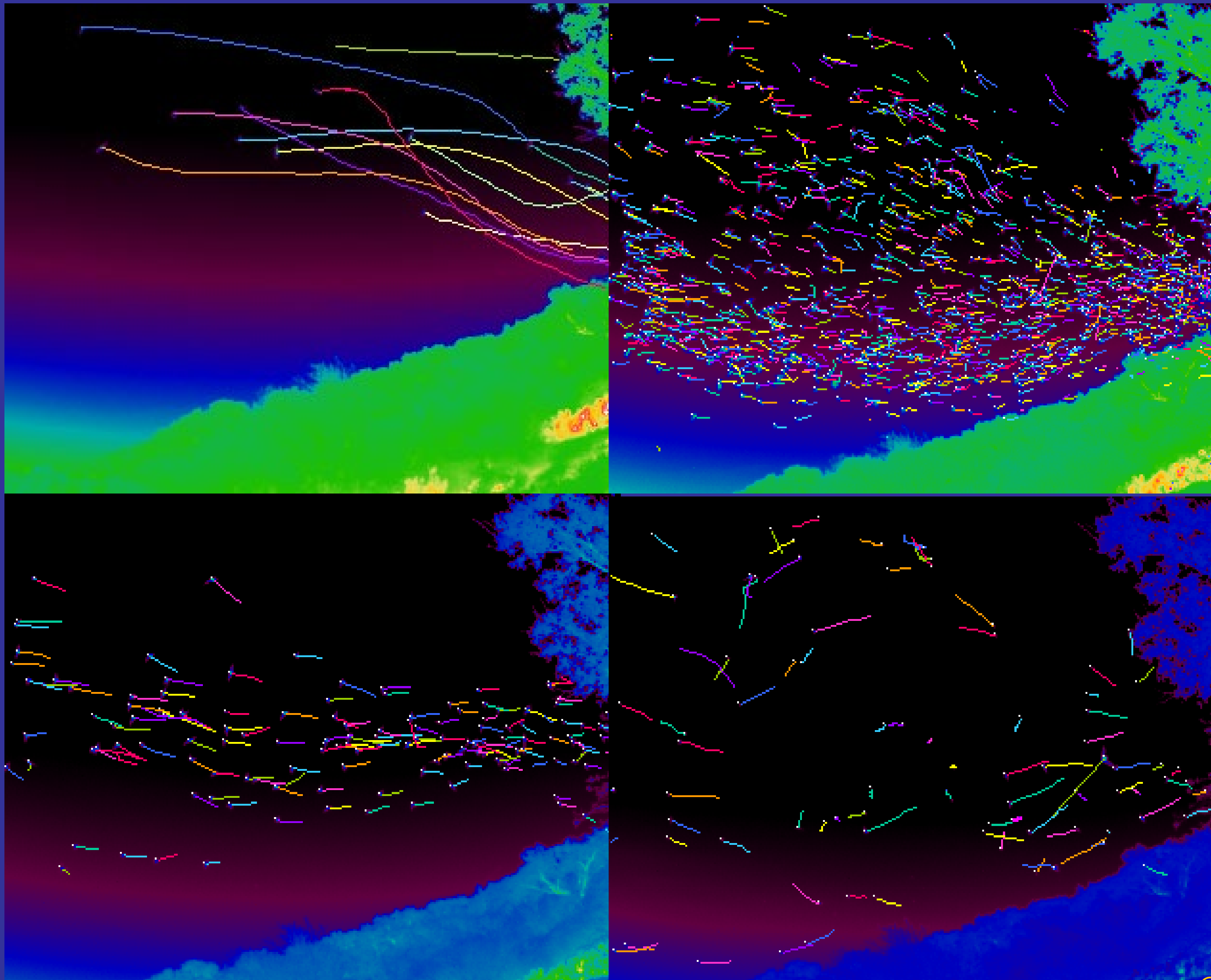
More on Kalman Filters next time

Tracking (Dis)Appearing Objects

- Occlusion:
  Window in time



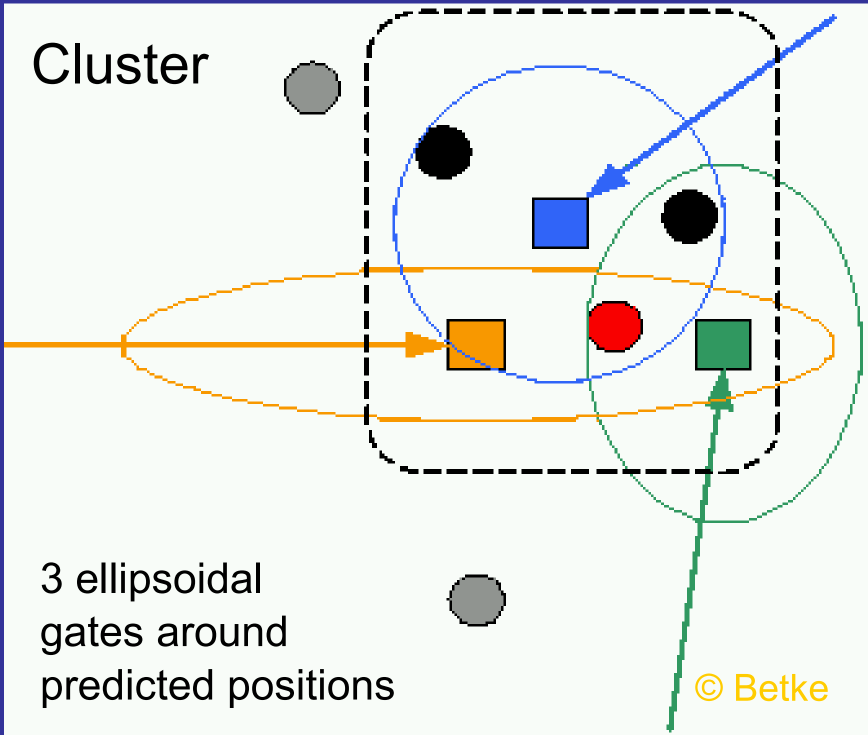

© Betke



Two Data Association Methods

- Both approaches based on gating: Prune number of candidate measurements so only measurements with likelihood within gate (= surface of constant probability density) must be considered.

- A cluster is created when the likelihood is high that a measurement (red disk) is due to any one of three objects (predicted positions shown as squares). Only measurements within cluster (black disks) are considered for assignment.



Poll: Which measurement should be assigned to which tracked object?

Predicted states: □
Measurements: ○

A:

Orange -> red

Green -> magenta

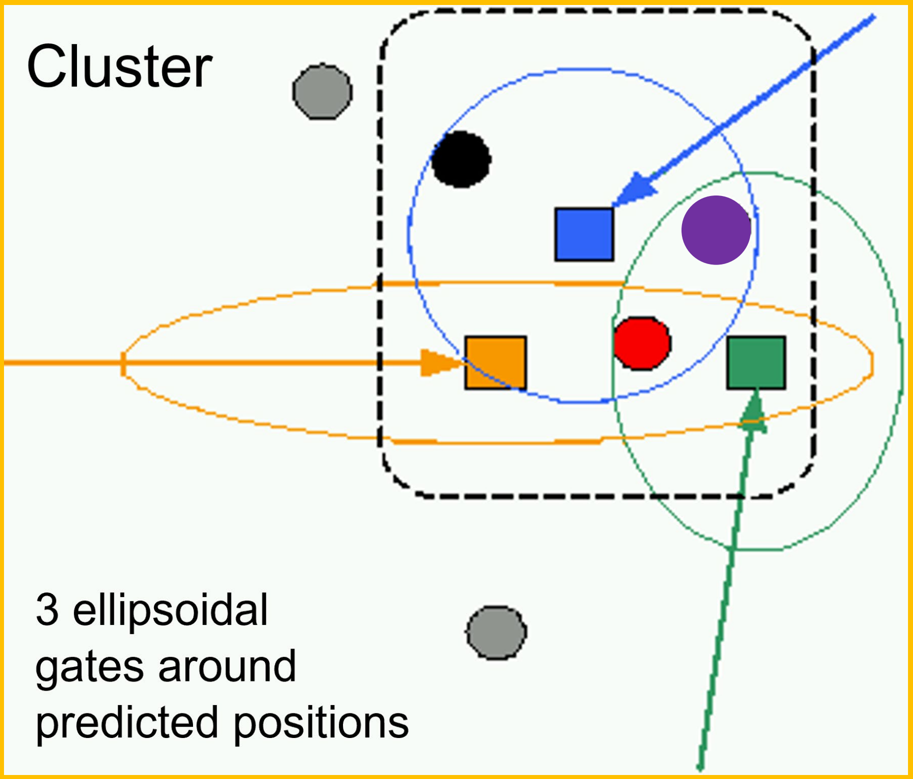
Blue -> black

B:

Blue -> red

Green -> magenta

Orange -> black



Poll: Which measurement should be assigned to which tracked object?

A:

Orange -> red

Green -> magenta

Blue -> black

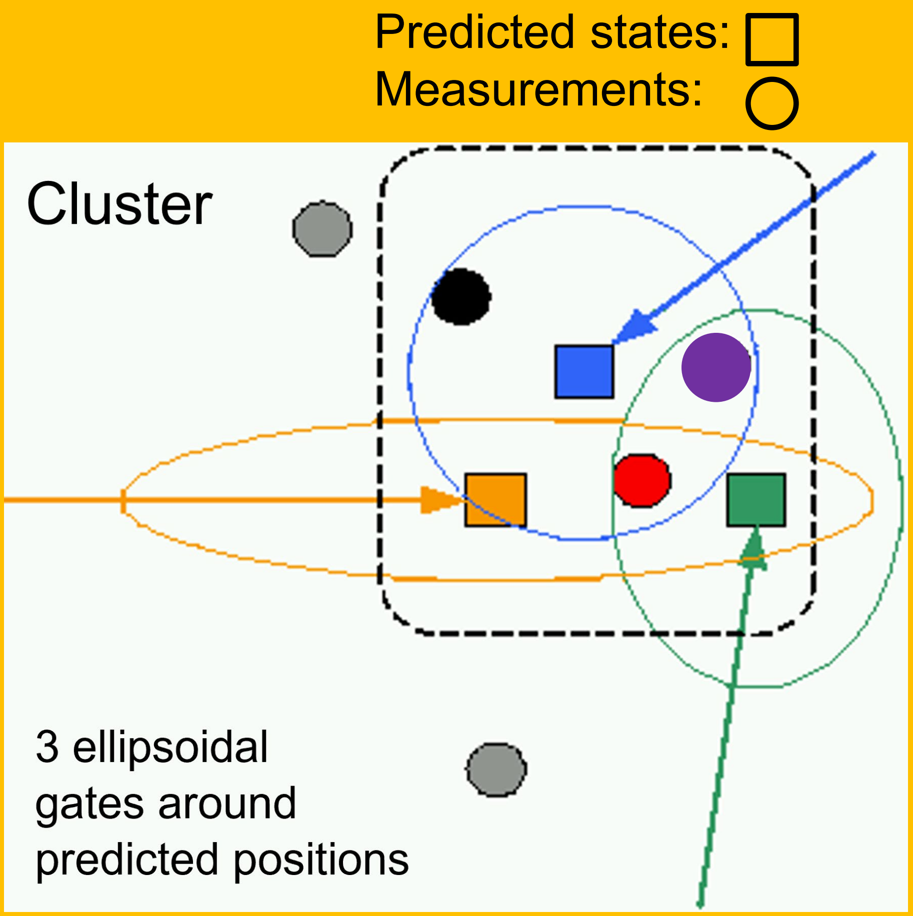
B:

Blue -> red

Green -> magenta

Orange -> black

The black measurement is not in the gate of the orange object.

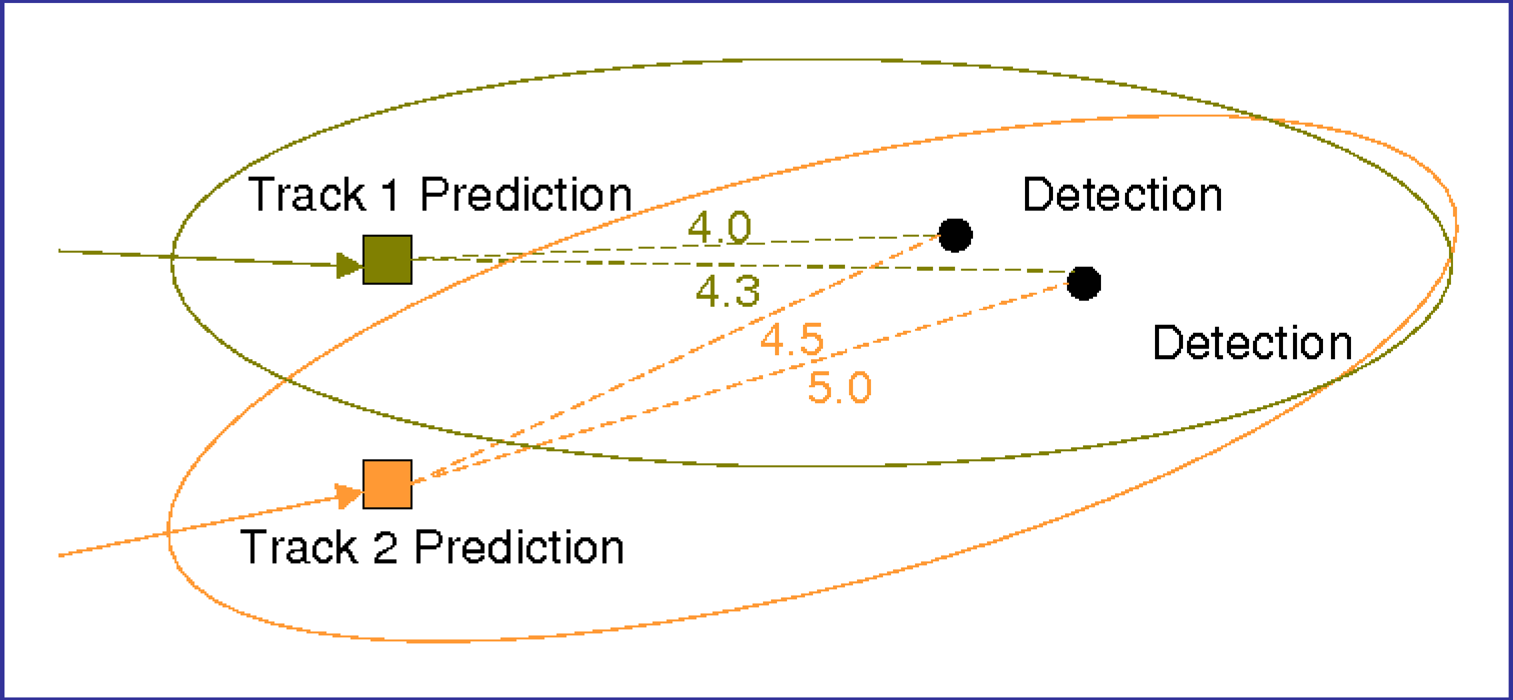


Two Data Association Methods

- Cluster-based Approach: uses Hungarian method to match measurements and objects in cluster.
- Greedy Approach: “greedily” favors objects with long observation histories. Matching process is started by matching longest-observed object and its nearest measurement. Then second-longest observed object in cluster is matched with its nearest measurement, etc.

Computation Complexity

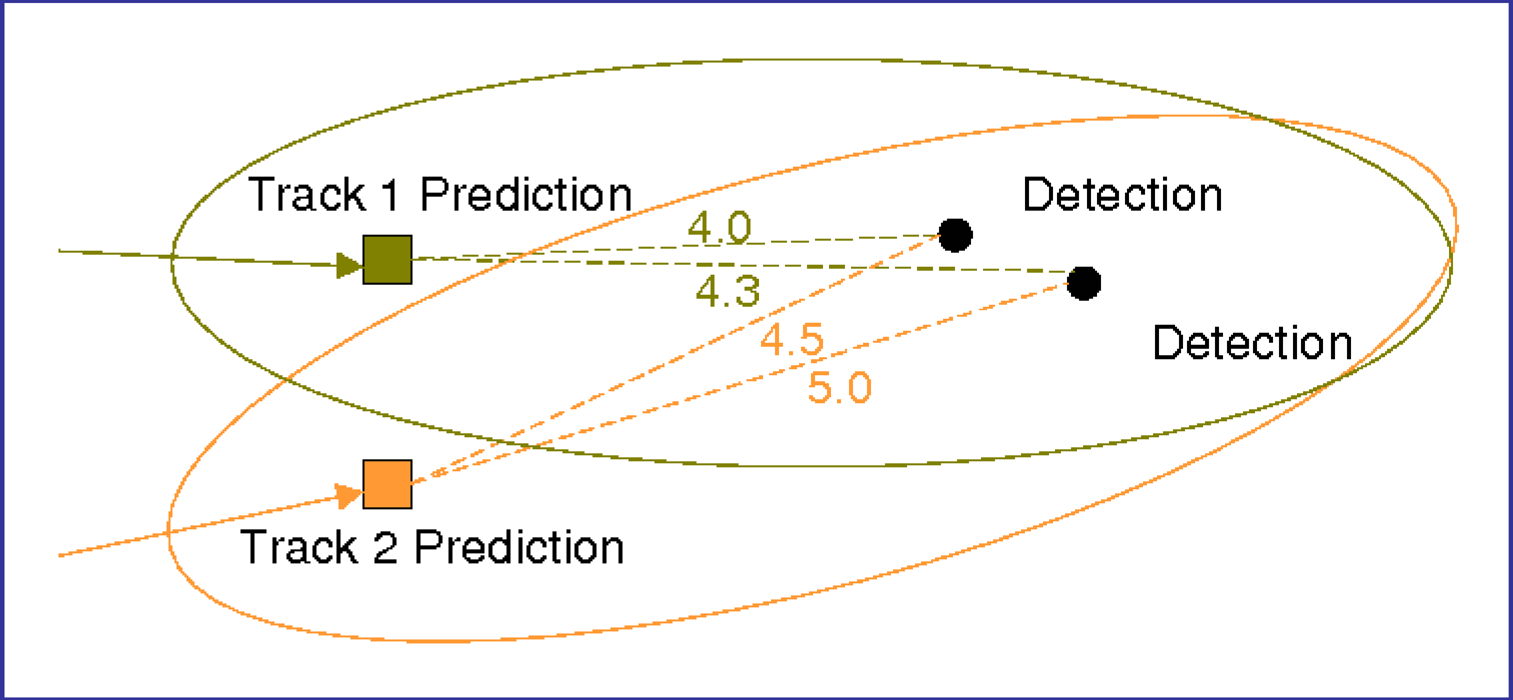
Data Association Methods



Greedy Method[1] selects measurement with min. distance (4.0) for "older" track 1. For "newer" track 2, measurement with distance 5.0 remains for assignment. Total distance in cluster is 9.0. $O(n^2)$, n = number of tracks

[1]Implemented by Diane Hirsh.

Data Association Methods



Greedy Method[1] selects measurement with min. distance (4.0) for "older" track 1. For "newer" track 2, measurement with distance 5.0 remains for assignment. Total distance in cluster is 9.0. $O(n^2)$, n = number of tracks

Hungarian Method[2] selects detections that minimize the total distance in the track cluster, here 8.8.
$O(n^3)$, n = number of tracks

[1]Implemented by Diane Hirsh.
[2]Implemented by Angshuman Bagchi

Censing Experiments

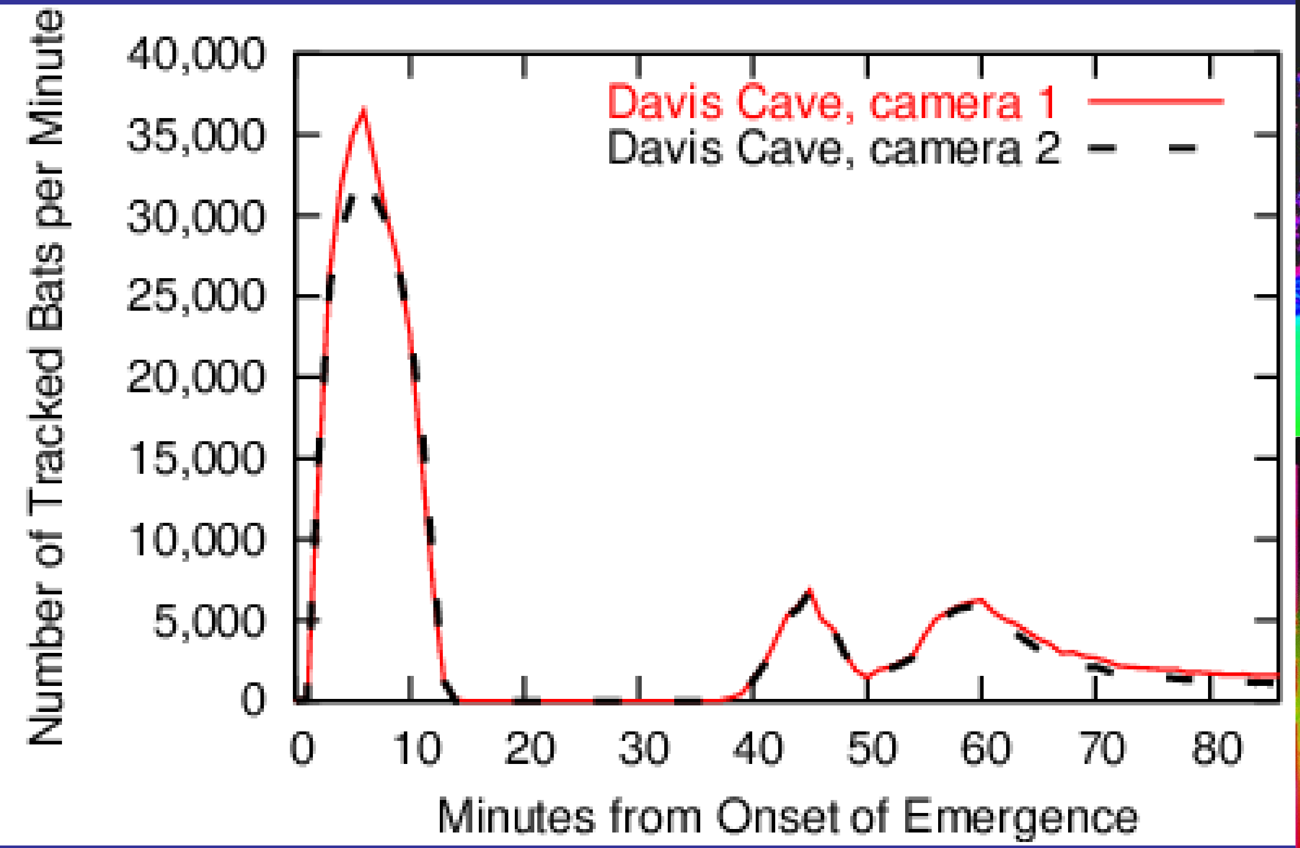
- Indigo Systems Merlin Mid Infrared Camera;
  ~60 Hz; 320 *x* 240 pixels of 12-bit intensities; infrared range: 1-5.4 μm.
- Processing rate was 10.8 Hz, now real time.
- 1st Validation experiment:
  2:32 min video (9,139 frames). 834,979 tracked objects; pruning with persistence threshold of 32 frames (can be interrupted by 5 frames of occlusion/low SNR);
  Ground truth: 7,007 bats. Our method: 7,056 bats (0.78% difference).

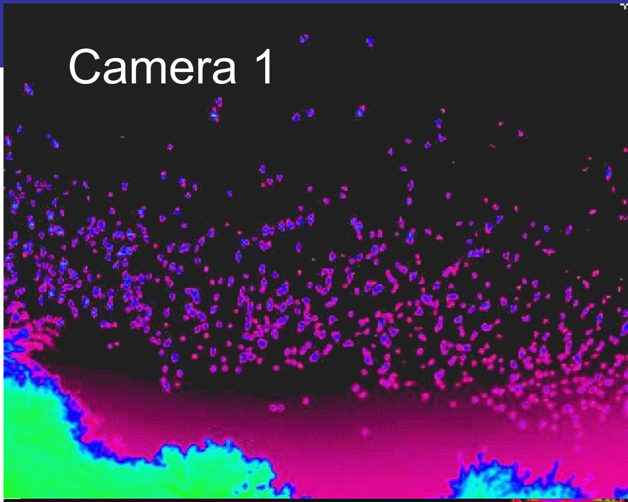
Tracking Emerging Bats

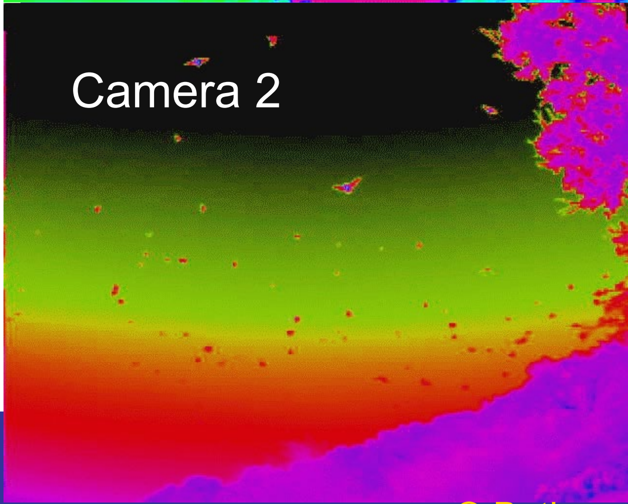
	Bamberger 07/03/2004	**Davis Cave** 07/16/2004
Processed frames	9,139	14,400
Analysis period	1 ½ minutes	4 minutes
Detected objects	834,979	8,743,240
Threshold for tracked object to count as a bat	0.5 second	0.5 second
Average # of live tracks	132.7	830.6
Tracked bats (a)	7,056	91,790
Manual estimate (m)	7,007	88,108
Deviation (a-m)/m	**0.7%**	**4.2%**

Another Validation Experiment

We compared results from two significantly different fields of views: 5.7% difference.

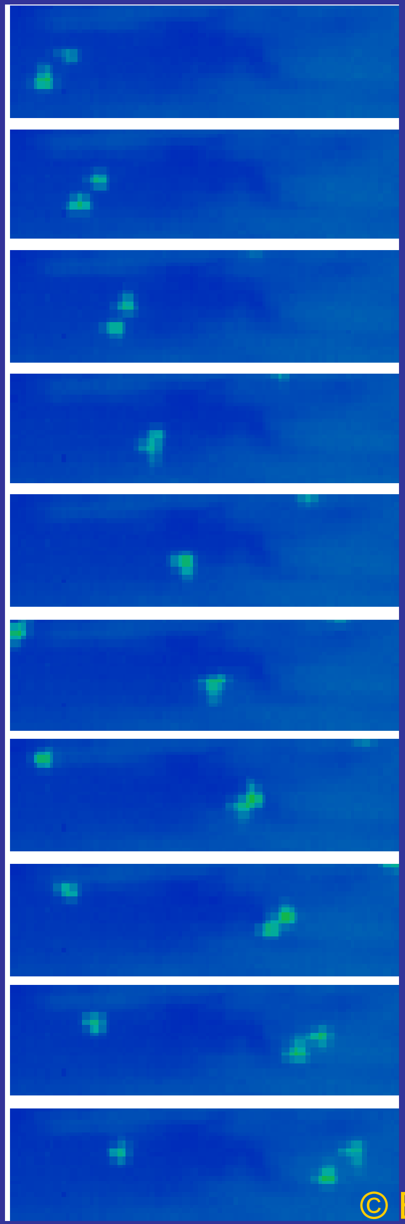


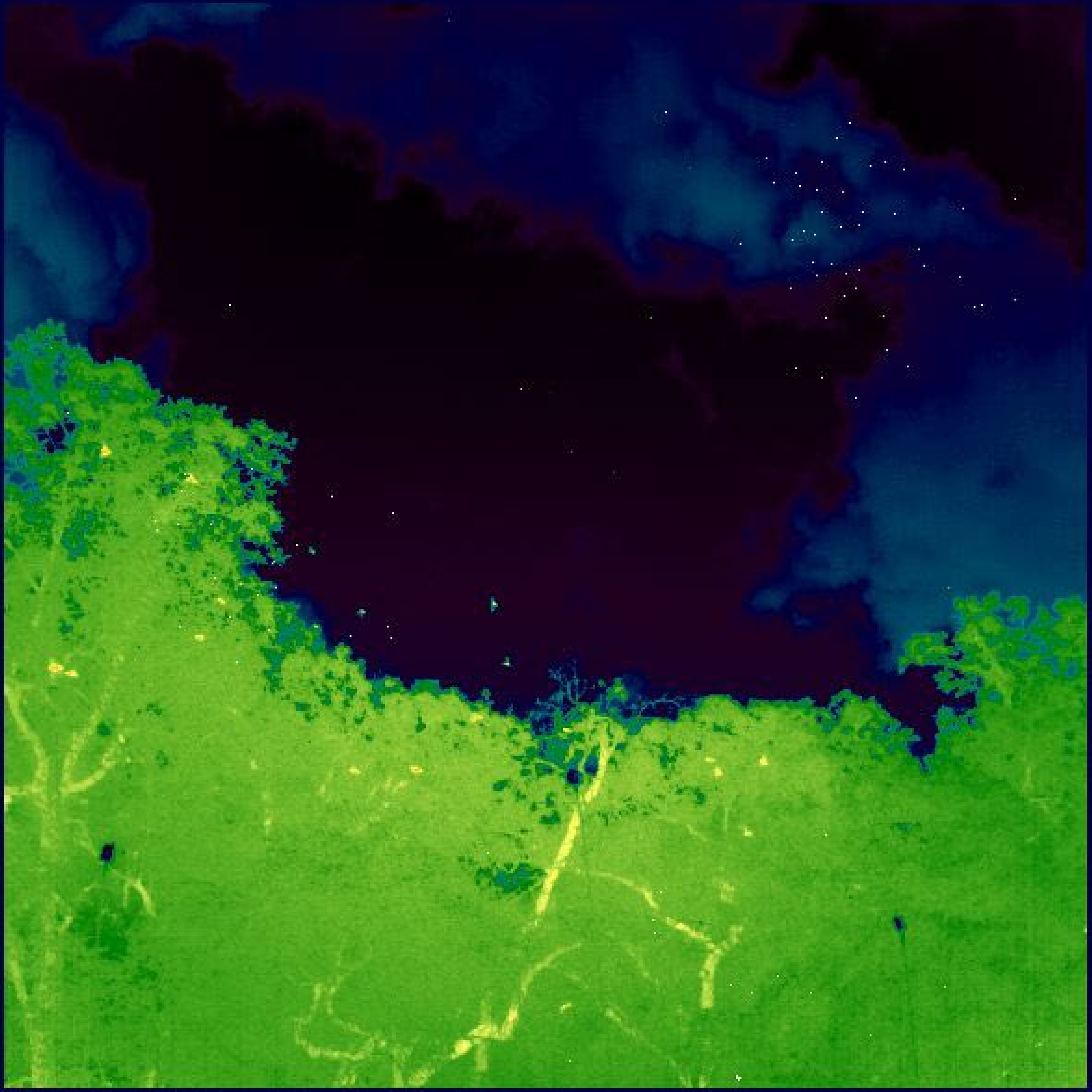


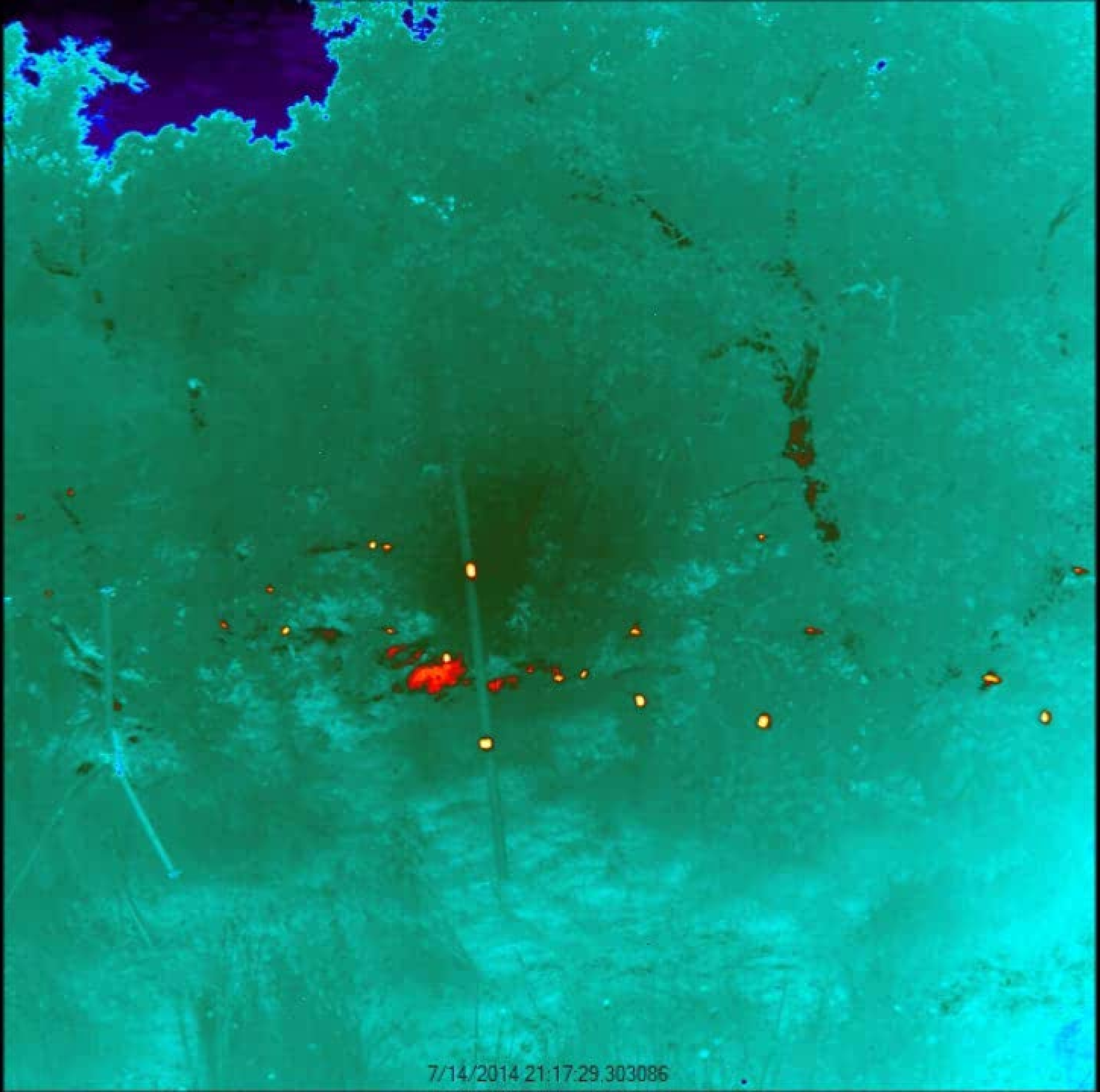


Conflict Resolution Experiments

- Decisions of cluster-based approach were similar to 4 independent volunteers (80% agreement vs. 20% for greedy approach).
- Ground truth is difficult to establish

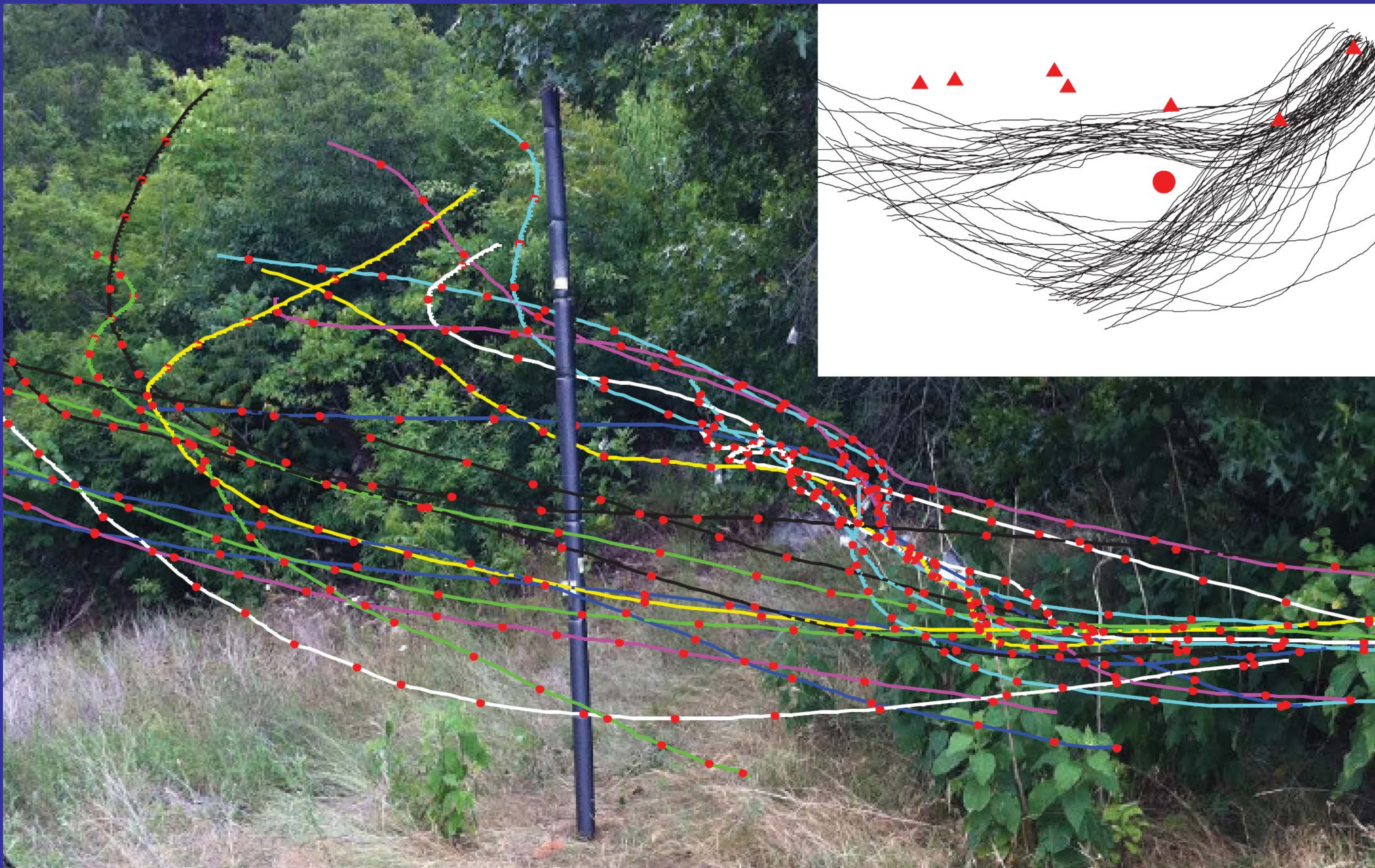





7/14/2014 21:17:29.303086

Opportunities for Studying Wildlife

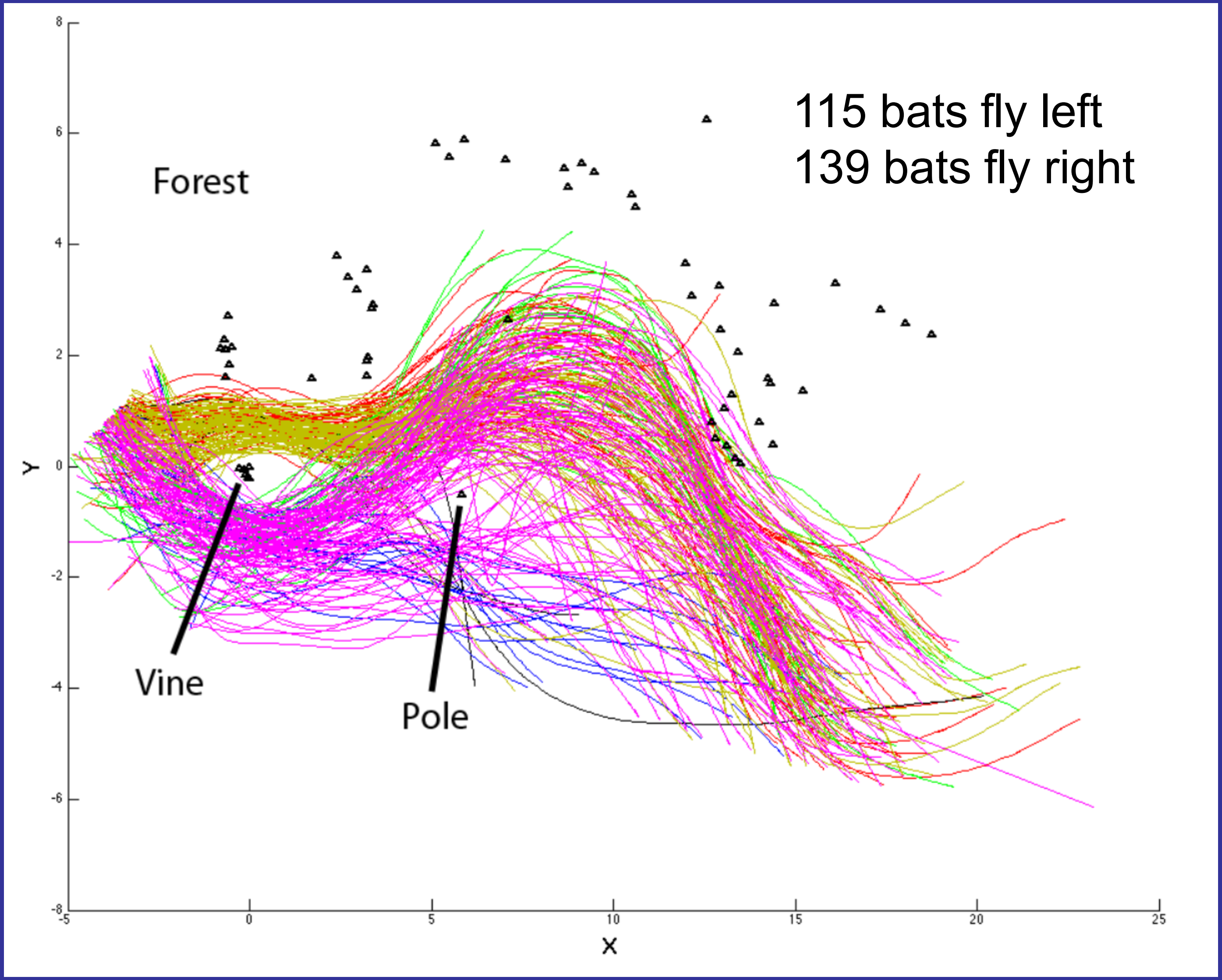
How do bats fly with respect to their environment?



Opportunities for Studying Wildlife

How do bats fly with respect to their environment?

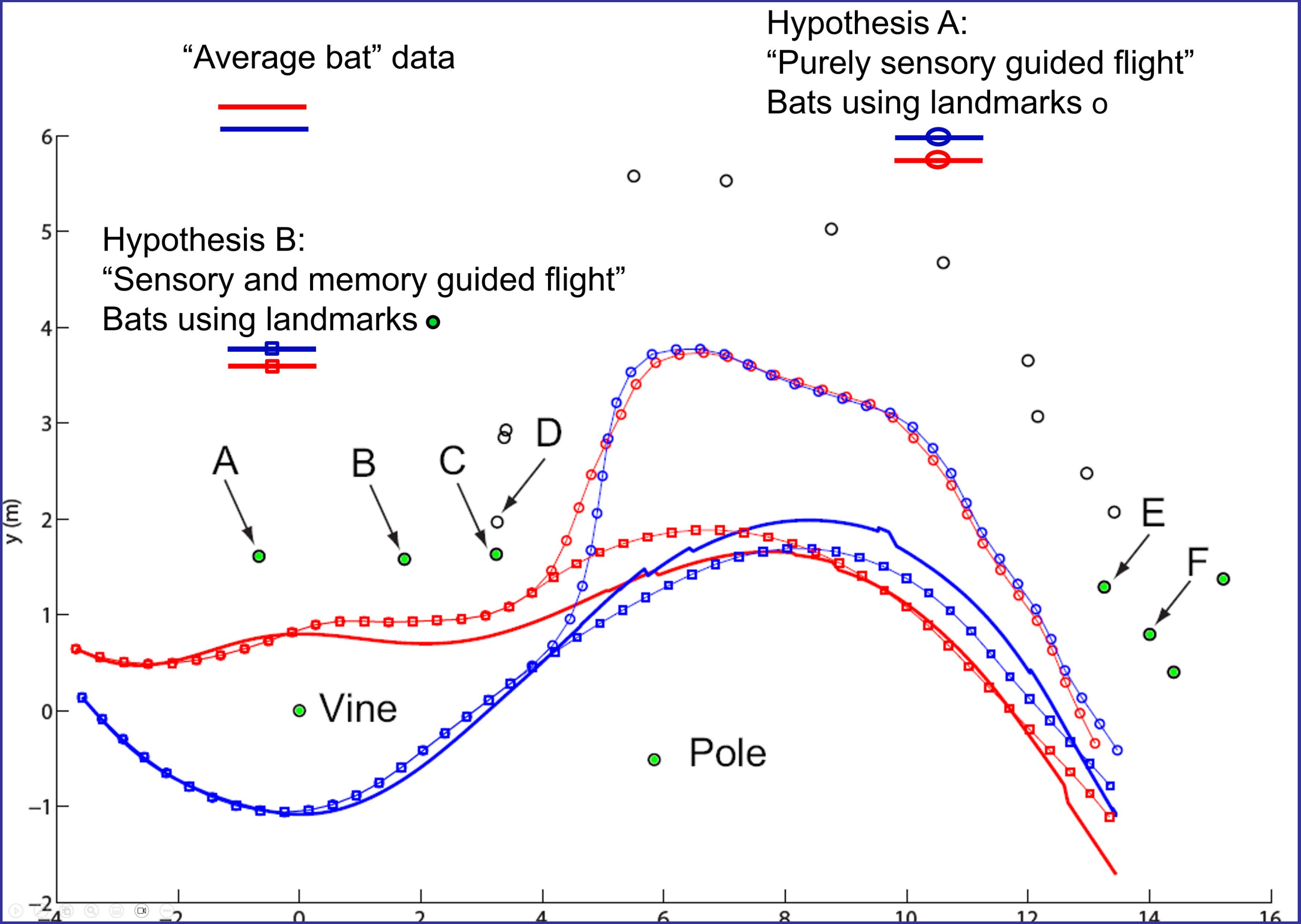
Cave



© Betke

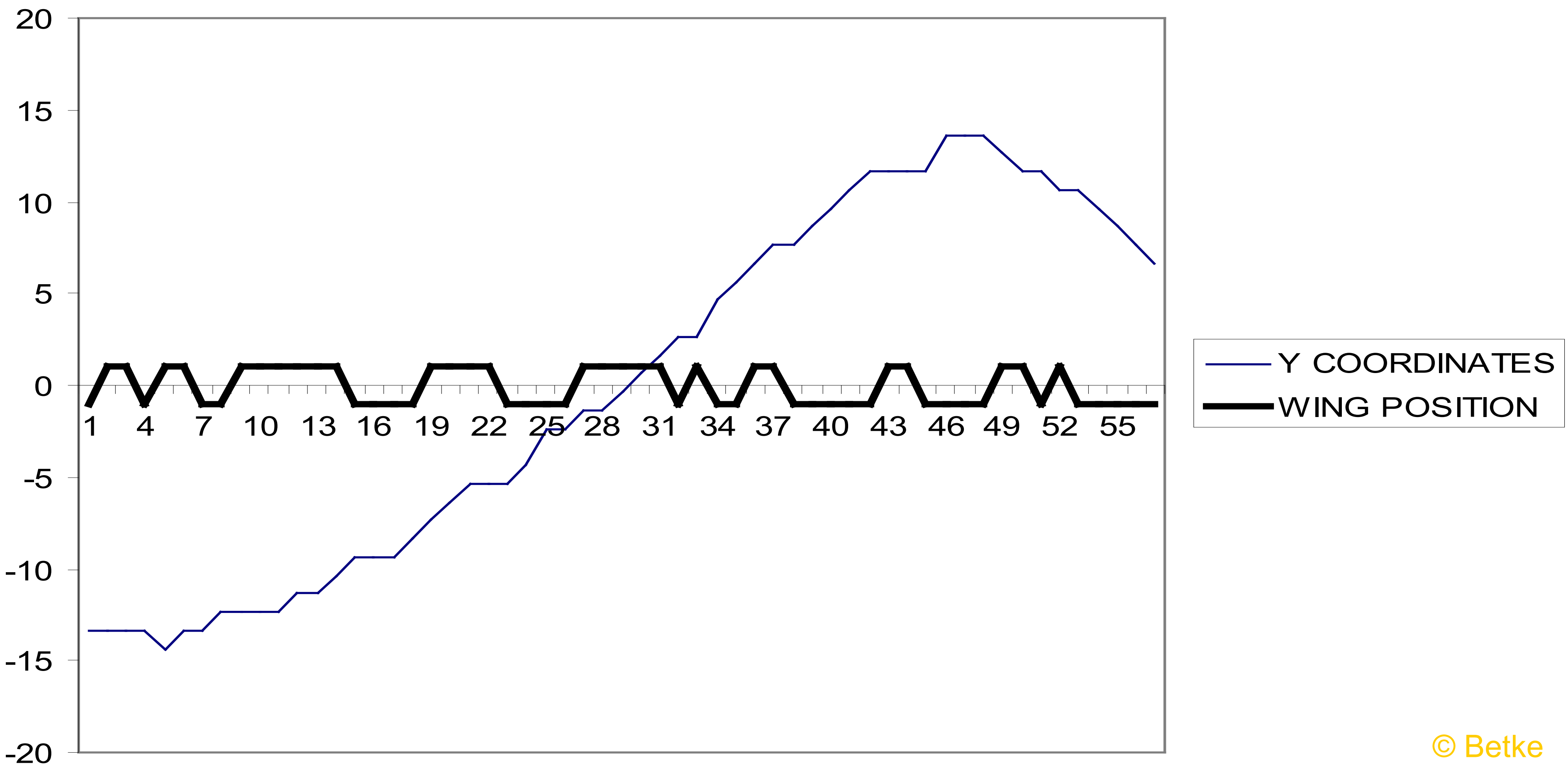
Opportunities for Studying Wildlife

How do bats fly with respect to their environment?

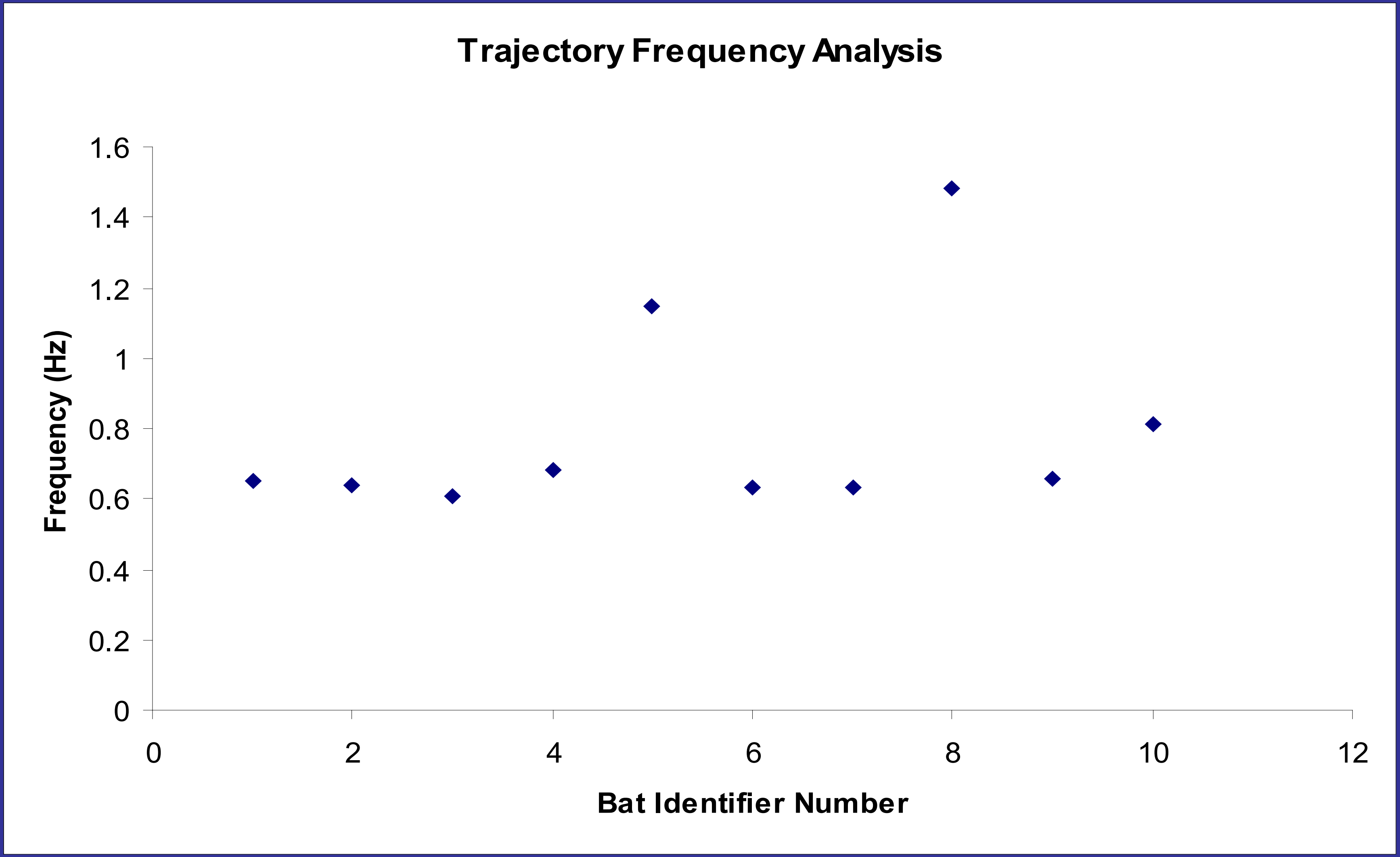


Analysis of Wing Beats

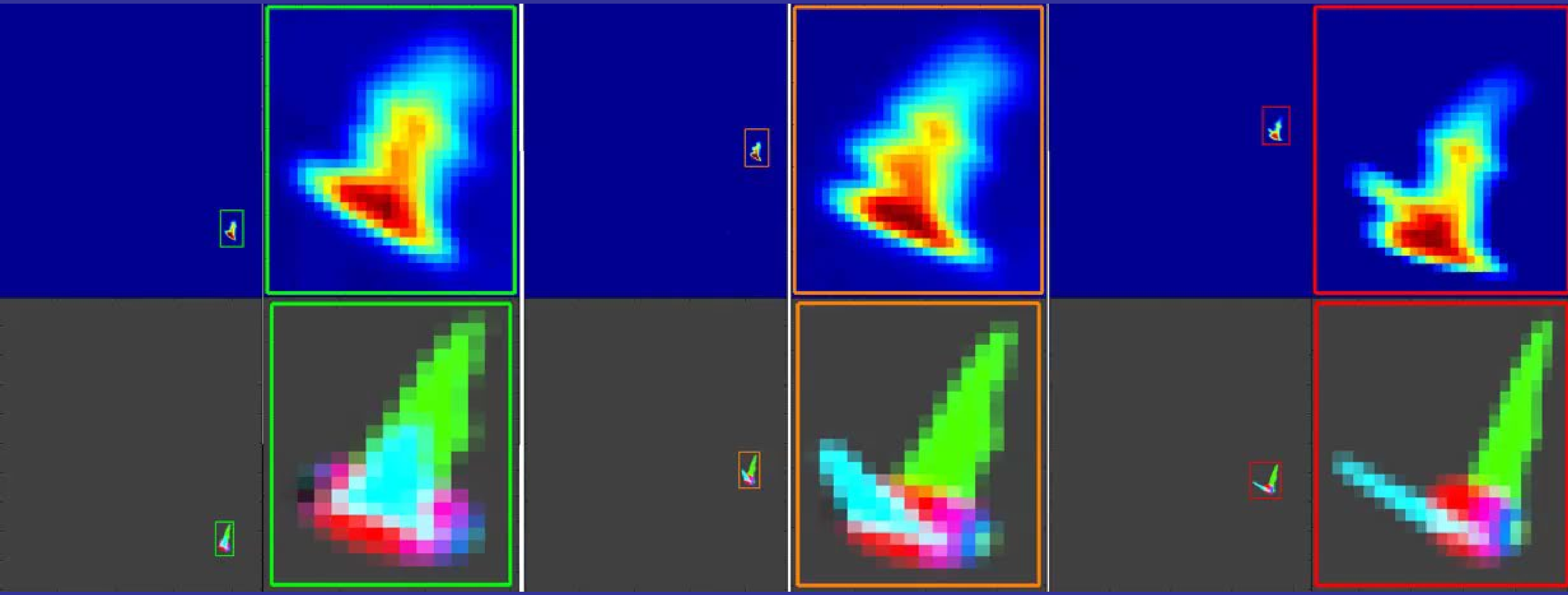
Wing beat frequency during ascent and descent of emerging bat



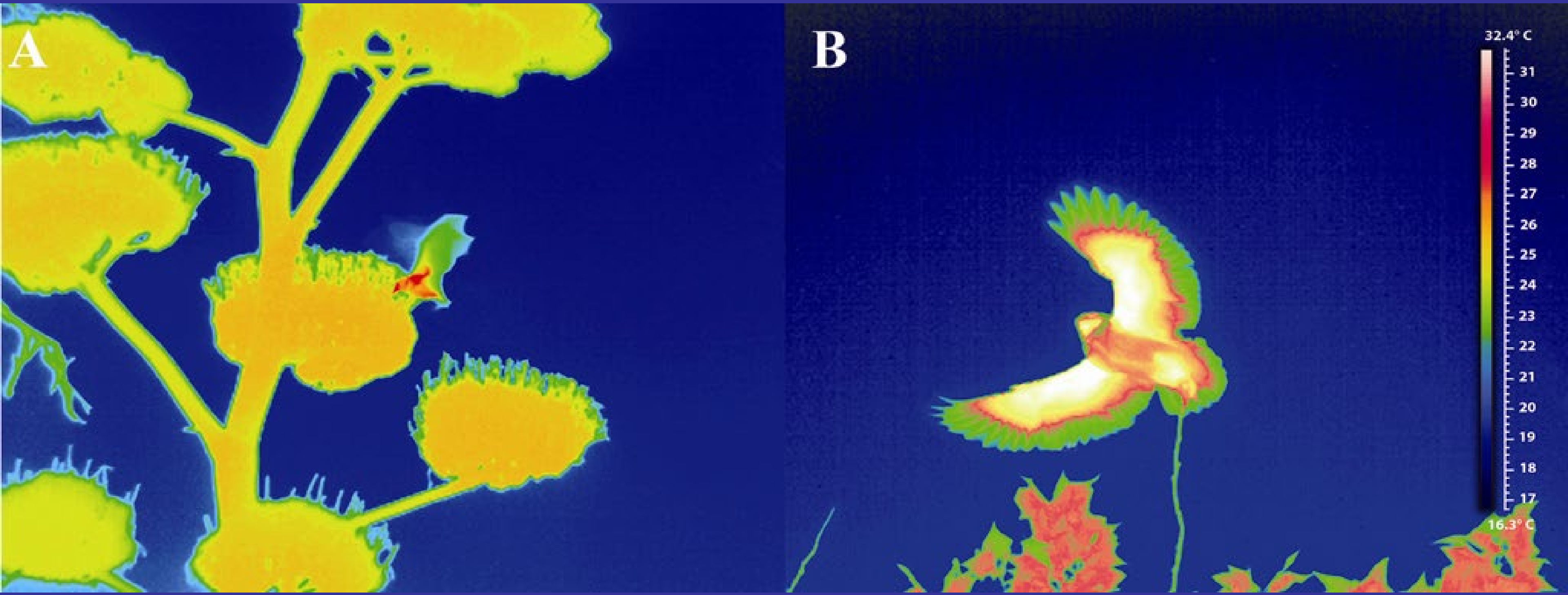
Frequency Analysis



Mean trajectory frequency of 0.8 Hz



- 3D model of bat
- Simulation of 2D projections of 3D model
- Comparison with real data


A
B
32.4° C
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16.3° C


Visit sometime!
© Betke

Learning Outcomes: Be able to

- Explain why it is challenging to census and track bats with computer vision tools
- Explain how bats can be detected
- Explain how bats can be tracked
- Explain what a Bayesian recursive filter is
- Define state, measurement, and update equations for alpha-beta and Kalman filters
- Explain what “tracking by detection” means
- Explain two data association methods
- Provide a high-level design (flow chart) of a multi-object tracking system
- Discuss how tracking systems can be validated experimentally