MiShape: Accurate Human Silhouettes and Body Joints from Commodity Millimeter-Wave Devices

Aakriti Adhikari, Hem Regmi, Sanjib Sur, Srihari Nelakuditi

Contact: aakriti@email.sc.edu
Monitoring At-home Human Activities

At-home monitoring unlocks numerous healthcare applications
Existing Techniques

Limited by Low Light and Low Visibility

Additional Hardware Requirements

Expensive and Privacy-invasive

Requires Extra Hardware

Provides Limited Information

Low Resolution in WiFi
Millimeter-Wave in 5G as an Opportunity
5G and MmWave

- Smaller Wavelength (mm)
- Illuminates Target Scene with High-Frequency
- Works Through Obclusion
- Robust to Low Light and Low Visibility
MiShape

Actual Activity

What vision camera produces

What MiShape produces

2.2m

5G
MiShape captures accurate shape information comparable to vision camera.
MiShape Overview

Human Performs Activity

Reflecting Signals

Combines Reflection from Multiple Antennas

Deep Joint Estimator

3D Human Joints

High Quality Human Silhouette

Generative Adversarial Networks

8x256
Challenges

Specularity and Weak Reflectivity

Low Imaging Resolution

Image Aliasing
Challenges

- Specularity and Weak Reflectivity
- Low Imaging Resolution
- Image Aliasing
Reflections From Human Body

Most of the signals transmitted do not reach back to the mmWave receiver
Reflections From Human Body

RGB Image

Depth Image

mmWave Image

Imperceptible human shape with many missing parts
Challenges

Specularity and Weak Reflectivity

Low Imaging Resolution

Image Aliasing
The resolution of mmWave images depends on the antenna array size.
Challenges

- Specularity and Weak Reflectivity
- Low Imaging Resolution
- Image Aliasing
Non-Uniform Antenna Placements

Aliasing due to non-uniform antenna placements in COTS mmWave devices
Non-Uniform Antenna Placements

Depth Image

MmWave Image
Leverage Signatures in Reflected Signals
Reflected Signal Analysis

There exists correlation between reflected signals and different postures
There exists correlation between reflected signals and different volunteers.
Generative Adversarial Networks

- Random Noise $z$ to Generator
- Fake Image $G(z)$ from Generator
- Real or Fake to Discriminator
- Discriminator to Real Data $(X)$
Conditional Generative Adversarial Networks

Random Noise $z$ \rightarrow \text{Generator} \rightarrow \text{Discriminator} \rightarrow \text{Real Data (X)}

Generator

Fake Image $G(z,y)$

Condition($y$)

Real or Fake

Real Data (X)
Conditional Generative Adversarial Networks
Conditional Generative Adversarial Networks

mmWave Reflections
8x256

Ground Truth

Conditional Generative Adversarial Networks (cGAN)

Epoch: 10

Imperceptible Image
Conditional Generative Adversarial Networks

mmWave Reflections
8x256

Ground Truth

Conditional Generative Adversarial Networks (cGAN)

Epoch: 100

Learning from Real Image Distribution
Conditional Generative Adversarial Networks (cGAN)

mmWave Reflections
8x256

Ground Truth

Conditional Generative Adversarial Networks (cGAN)

Epoch: 1000

Perceptible Human Silhouette
Conditional Generative Adversarial Networks

mmWave Reflections
8x256

Generator

Post Training

Perceptible Human Silhouette
Generating High-Resolution Silhouette

Super-Resolution GAN (SRGAN)

Improves resolution and recovers any missing information
Implementation

**mmWave Hardware**
- 77–81 GHz mmWave transceivers
- BW 1.62 GHz
- TI IWR1443BOOST
- Each with one transmit and four receive antennas

**Ground Truth**
- Microsoft Kinect Xbox One

Two transceivers with one rotated 90° counter-clockwise w.r.t. another
Data Collection

- Subject is asked to stand at approximately 2 m distance from the setup
- Dataset includes input-output pairs of mmWave reflections, human silhouette images, and 3D joint locations

Baseline Data Collection
- We collect datasets from a single subject with 17 diverse poses
- Each experiment takes 12 seconds to complete

Additional Data Collection
- We collect data from additional 9 volunteers for 5 diverse poses
- Each experiment takes 60 seconds to complete

In total, 100 K input-output pairs from 10 volunteers with diverse ages, gender, height, and somatotype
Specularity and Weak Reflectivity
Low Imaging Resolution
Image Aliasing
Evaluation

Specularity and Weak Reflectivity

Low Imaging Resolution

Image Aliasing
MiShape generates human perceptible complete silhouette
Full Silhouette Recovery with MiShape

MiShape’s cGAN model produces images close to ground truth
Evaluation

Specularity and Weak Reflectivity

Low Imaging Resolution

Image Aliasing
High-resolution Imaging

High-resolution silhouette generated by Generator similar to ground truth
Evaluation

Specularity and Weak Reflectivity

Low Imaging Resolution

Image Aliasing
MiShape is consistently better across diverse antenna configurations
Evaluation under Different Conditions

Works Under Occlusion

Works Under Low Light and Low Visibility

Works in Presence of Multiple Objects
Gait Monitoring

MiShape follows ground truth walking trajectories well.
MiShape generates high-quality human silhouettes and predicts 3D locations of body joints on par with existing vision-based systems.

MiShape brings high-resolution, through-occlusion imaging into ubiquitous commodity 5G devices.

MiShape enables application in gait monitoring with higher accuracy.

Thank you!

Please check out our paper for more results:

Any Questions: Please email to aakriti@email.sc.edu