Knowledge Base Creation by Reliability of Coordinates Detected from Videos for Finger Character Recognition

Shinpei Hagimoto*, Takuma Nitta*, Ari Yanase*,
Takafumi Nakanishi*, Ryotaro Okada*, Virach Sornlertlamvanich*
Department of Data Science Musashino University*

19th IADIS International Conference e-Society 2021
05/03/2021 08:05-10:15 Session FSP 5.1 _ eLearning // Knowledge Management _ F144

Musashino University

Outline

- Background
- Purpose
- Proposed method
- Experiment and Evaluation
- Conclusion

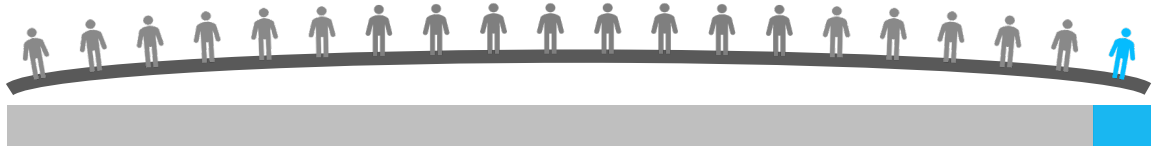
Outline

- Background
- Purpose
- Proposed method
- Experiment and Evaluation
- Conclusion

Background

Social

A sign language translation system is very significant.



About 5% of the world population has hearing loss

Academic

Not enough datasets

Using special device for data extraction is costly

MLDL require a huge amount of image datasets and high computational cost

Background

Social

A sign language transition system is very significant.

About 5% of the world population has hearing loss

Academic

Not enough datasets

Using special device for data extraction is costly

MLDL require a huge amount of image datasets and high computational cost

Nitta, T., et al, 2020. Finger Character Recognition in Sign Language Using Finger Feature Knowledge Base for Similarity Measure. *3rd IEEE/ IIAI International Congress on Applied Information Technology*. Tokyo, Japan.

Background

Social

A sign language transition system is very significant.

About 5% of the world population has hearing loss

Academic

Not enough datasets

Using special device for data extraction is costly

MLDL require a huge amount of image datasets and high computational cost

Nitta, T., et al, 2020. Finger Character Recognition in Sign Language Using Finger Feature Knowledge Base for Similarity Measure. *3rd IEEE/ IIAI International Congress on Applied Information Technology*. Tokyo, Japan.

Background

Social

A sign language transition system is very significant.

About 5% of the world population has hearing loss

Academic

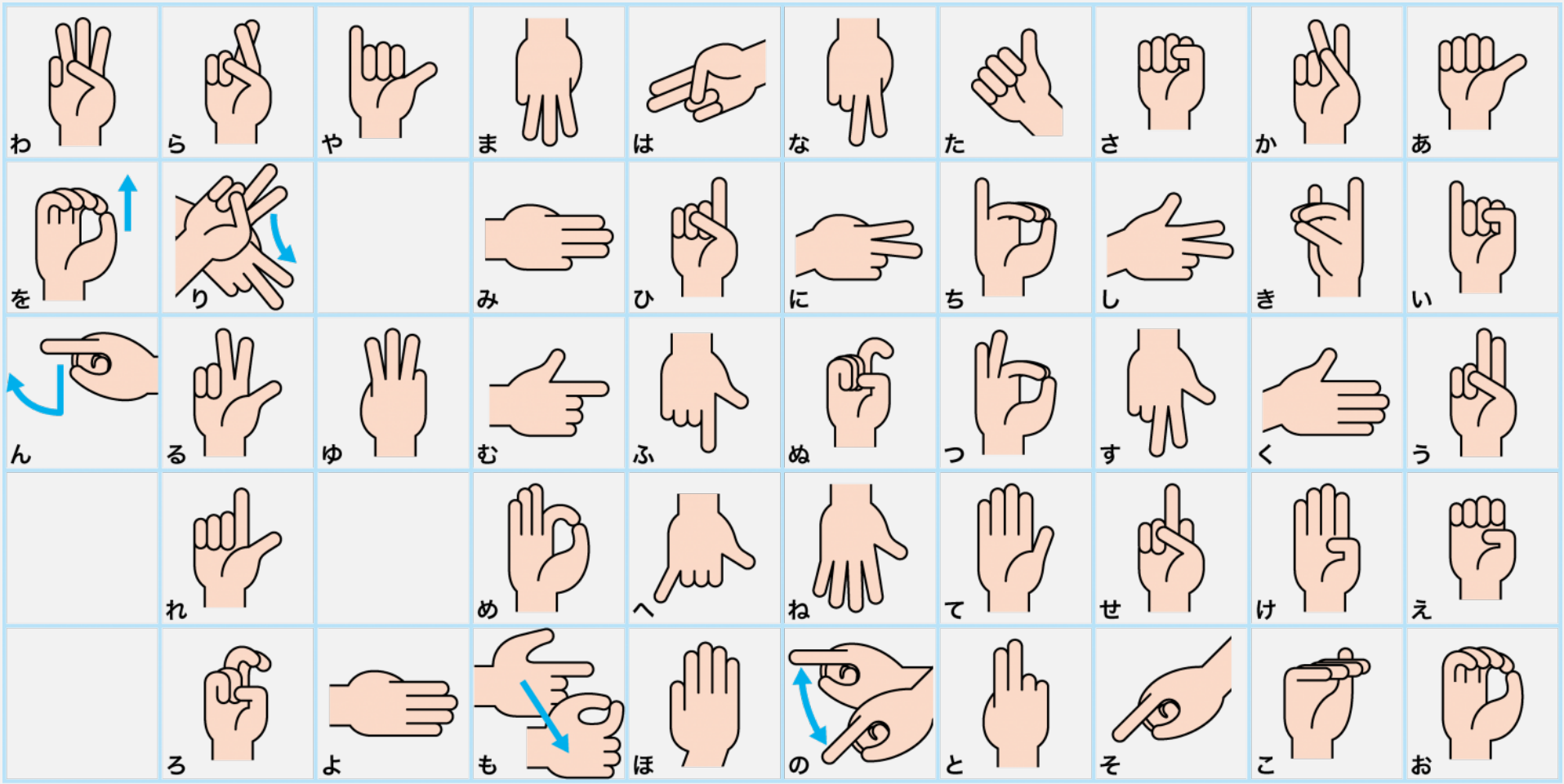
Not enough datasets

Using special device for data extraction is costly

MLDL require a huge amount of image datasets and high computational cost

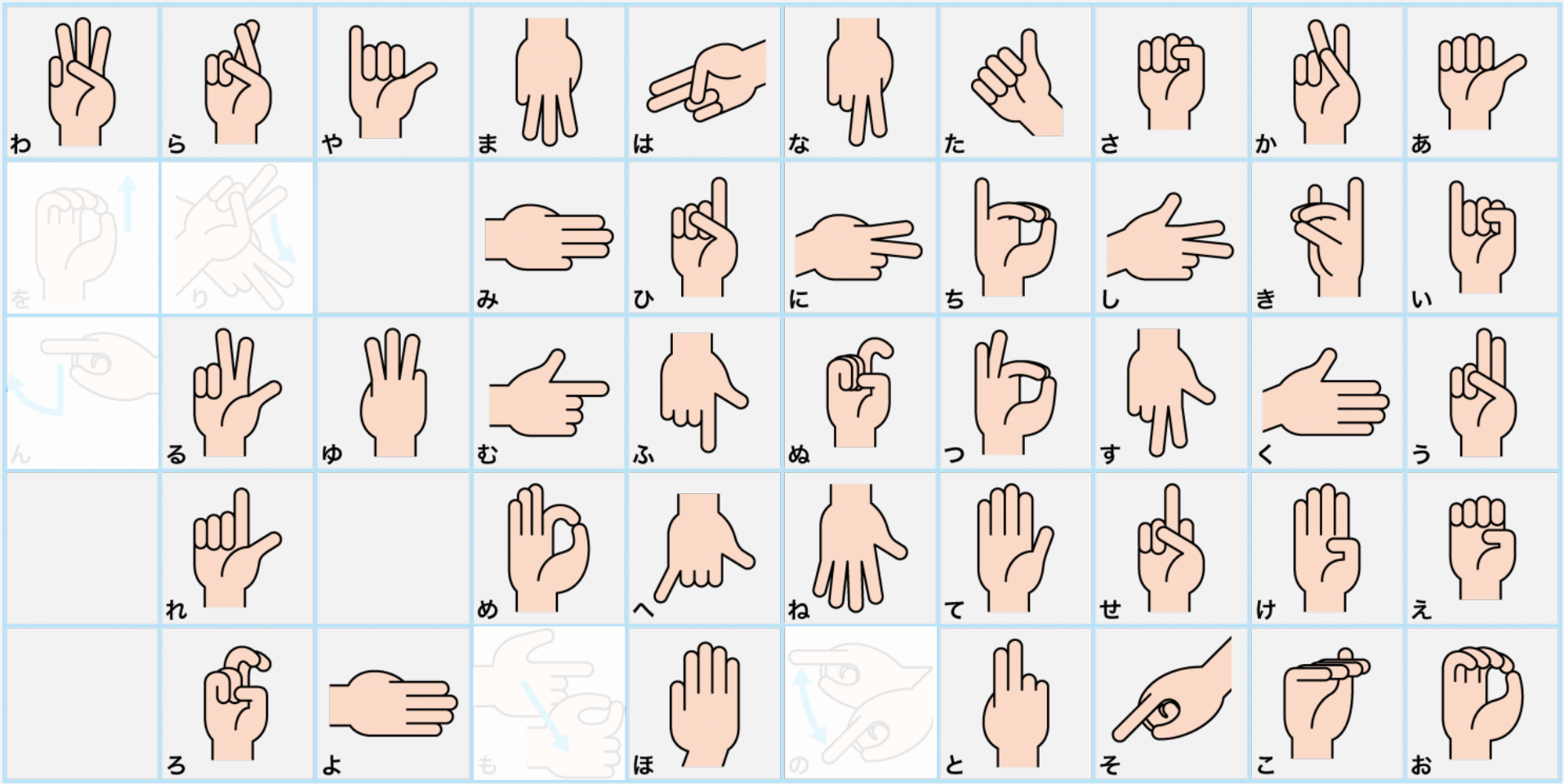
Nitta, T., et al, 2020. Finger Character Recognition in Sign Language Using Finger Feature Knowledge Base for Similarity Measure. *3rd IEEE/ IIAI International Congress on Applied Information Technology*. Tokyo, Japan.

Japanese finger character



ILLUST BOX https://www.illust-box.jp/sozai/101472/ (2021/3/2)

Japanese finger character



ILLUST BOX https://www.illust-box.jp/sozai/101472/ (2021/3/2)

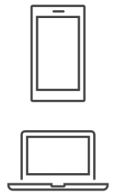
Outline

- Background
- Purpose
- Proposed method
- Experiment and Evaluation
- Conclusion

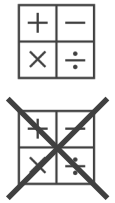
Main purpose

To construct a finger character recognition system

with a small image data set



to use a generic device



with low computational cost

Nitta, T., et al, 2020. Finger Character Recognition in Sign Language Using Finger Feature Knowledge Base for Similarity Measure. *3rd IEEE/ IIAI International Congress on Applied Information Technology*. Tokyo, Japan.

Purpose of this work

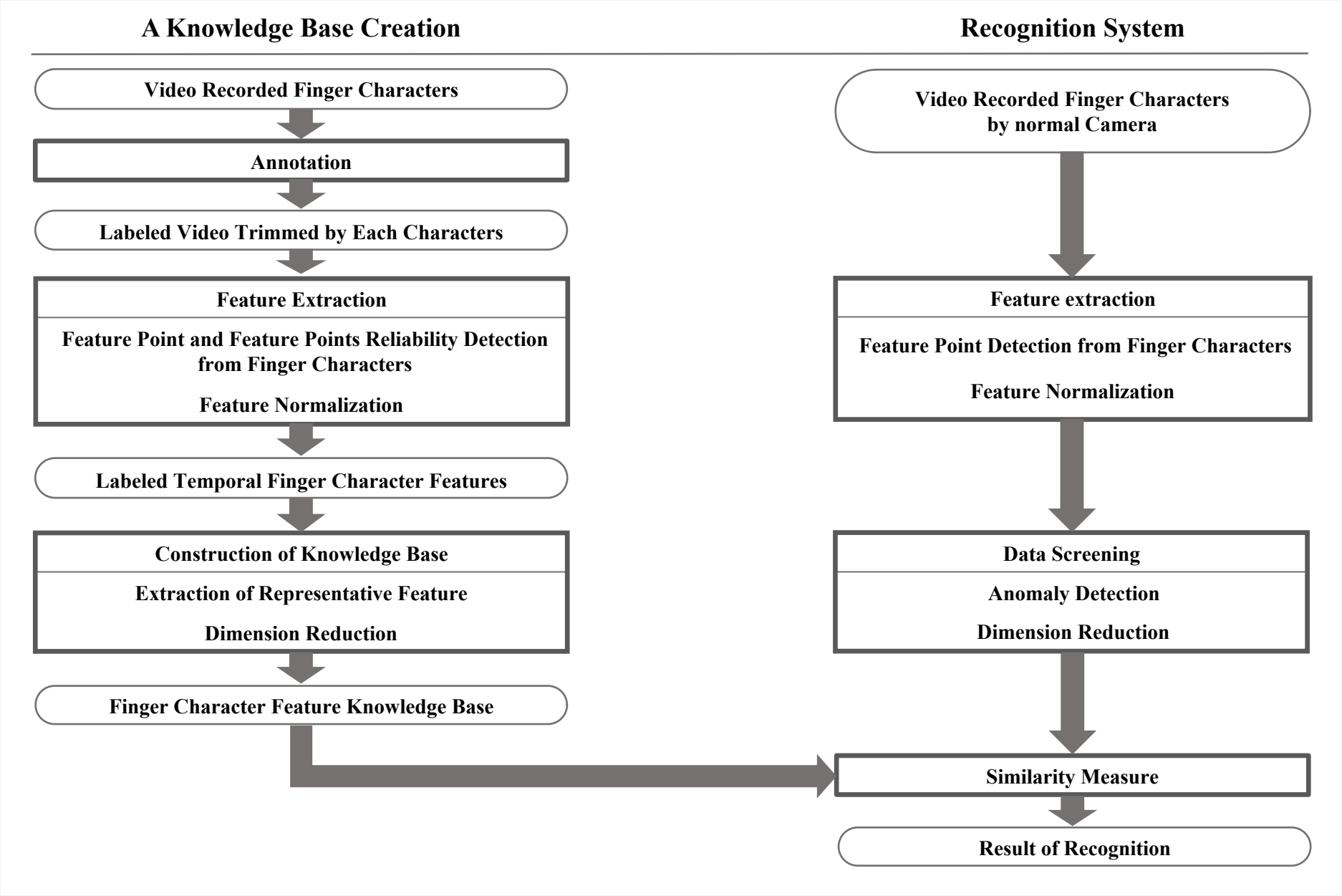
- We propose new knowledge base creation methods using the reliability of the finger joint coordinates.
- We evaluate the knowledge base creation method.
- We implement a finger character recognition system with higher recognition accuracy.

Outline

- Background
- Purpose
- Proposed method
- Experiment and Evaluation
- Conclusion

Overview of our recognition system

Proposed method



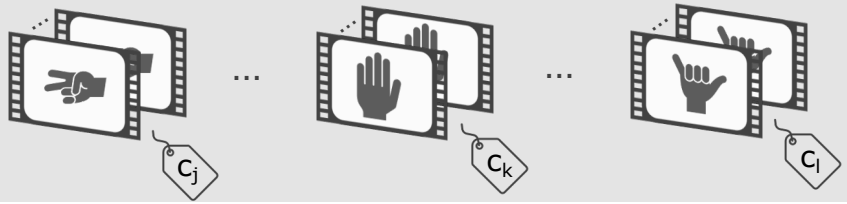
A knowledge base creation system

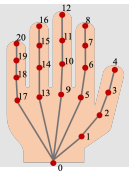
Proposed method

A Knowledge Base Creation

Finger Character Video

Annotation





Feature Extraction using OpenPose

Feature Normalization

Labeled Temporal Finger Character Features

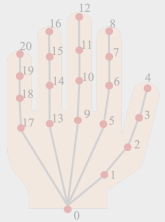
Representative Feature Dimension Reduction

Knowledge Base

Recognition System

Videos and images by normal camera

Feature Extraction using OpenPose



Feature Normalization

Data Screening

Anomaly Detection
Dimension Reduction

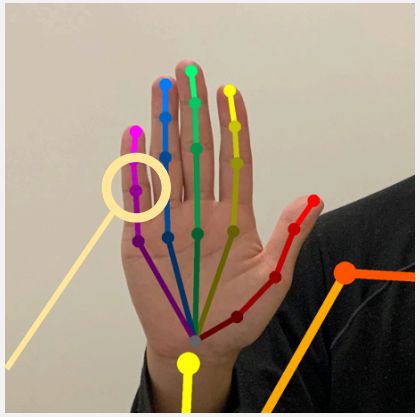
Similarity Measure

Result of Recognition

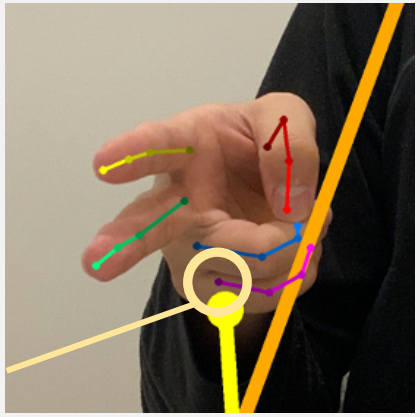
Definition of the reliability of coordinates

The reliability is a value in the range 0~1 that shows accuracy of the coordinates obtained by OpenPose.

(x, y, 0.8946)



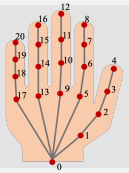
(x, y, 0.5881)



(X-coordinate, Y-coordinate, the reliability of the coordinates)

Improvement point

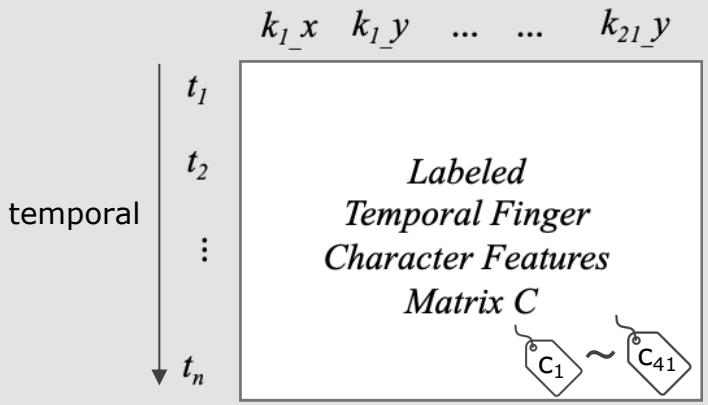
Proposed method



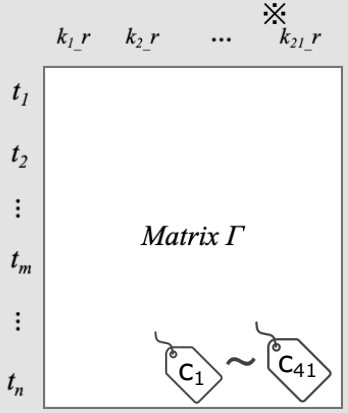
Feature Extraction using OpenPose

Finger joint coordinates

the reliability of coordinates



※



Labeled Temporal Finger Character Features

New Knowledge Base Creation Methods

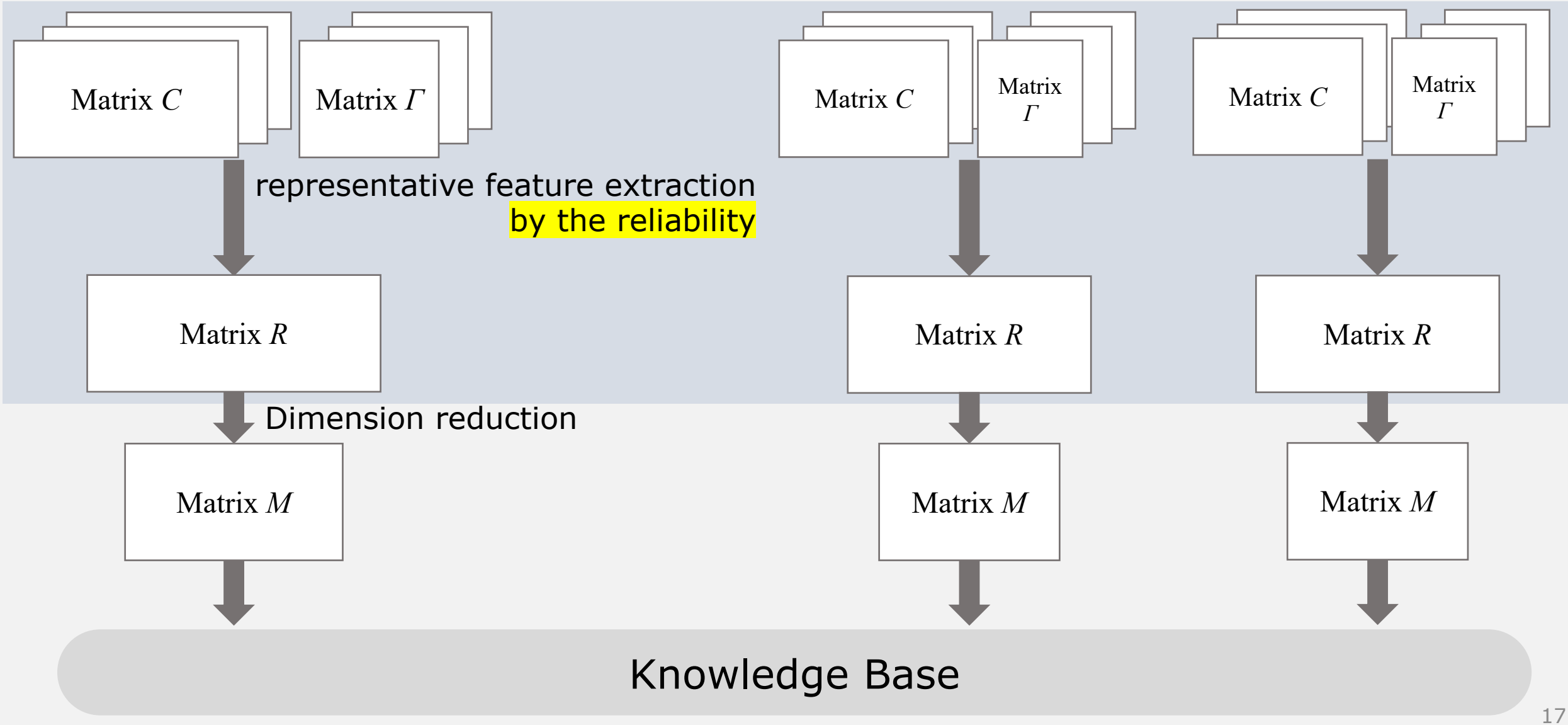
new knowledge base creation methods by reliability of coordinates

Knowledge Base

※ k_i_x : The coordinate of X k_j_y : The coordinate of Y k_l_r : The reliability of the coordinates (details in next slide)

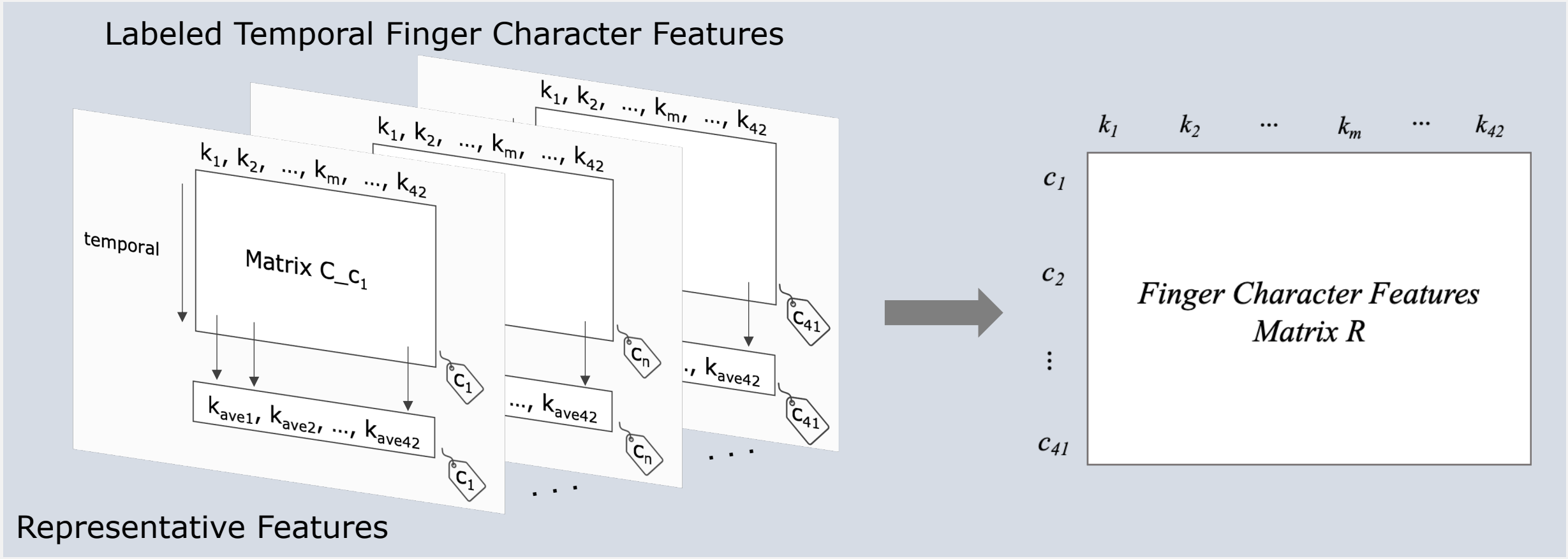
Outline of new knowledge base creation methods by reliability

Proposed method



Representative feature extraction

Proposed method



$k_{1\text{-}42}$: X and Y coordinates for each of the 21 finger joint coordinates

$c_{1\text{-}n}$: Each finger character

Nitta, T., et al, 2020. Finger Character Recognition in Sign Language Using Finger Feature Knowledge Base for Similarity Measure. *3rd IEEE/ IIAI International Congress on Applied Information Technology*. Tokyo, Japan.

Representative feature extraction

1. Creation by **the highest average value**
2. Creation by **the highest minimum value**
3. Creation by **the average of averages**
4. Creation by **the sum of the maximum and minimum values**

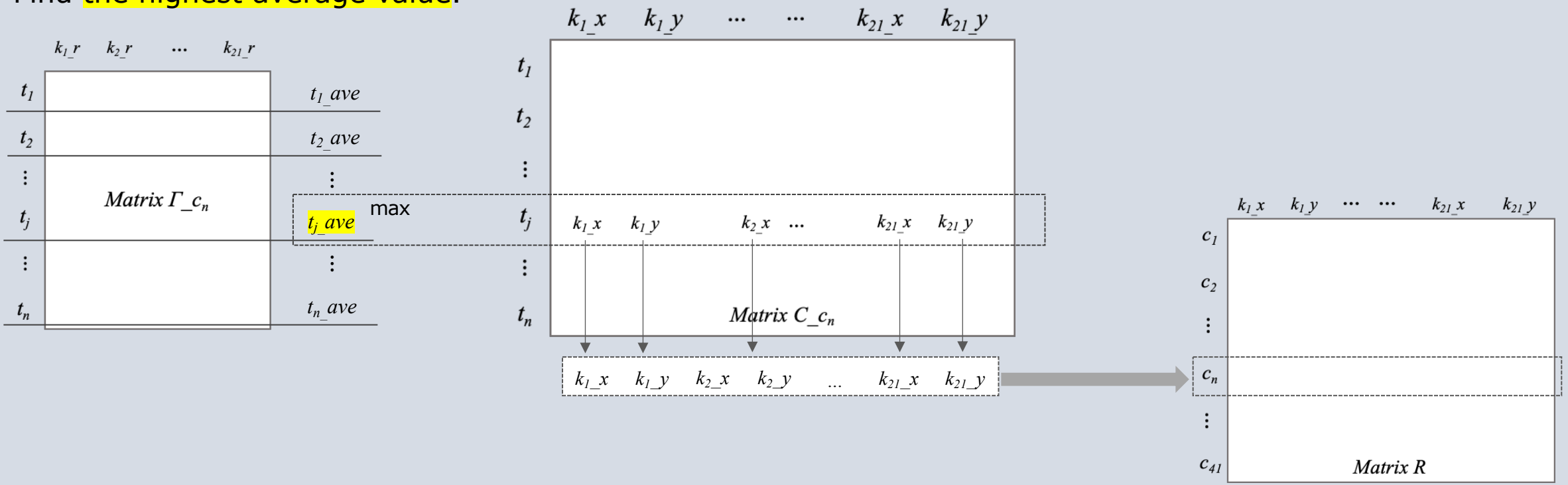
① Creation by the highest average value

For each frame, calculate the average of the reliability.

Find the highest average value.

Extract the coordinates of the frame with the highest average.

Store the extracted representative values in the knowledge base



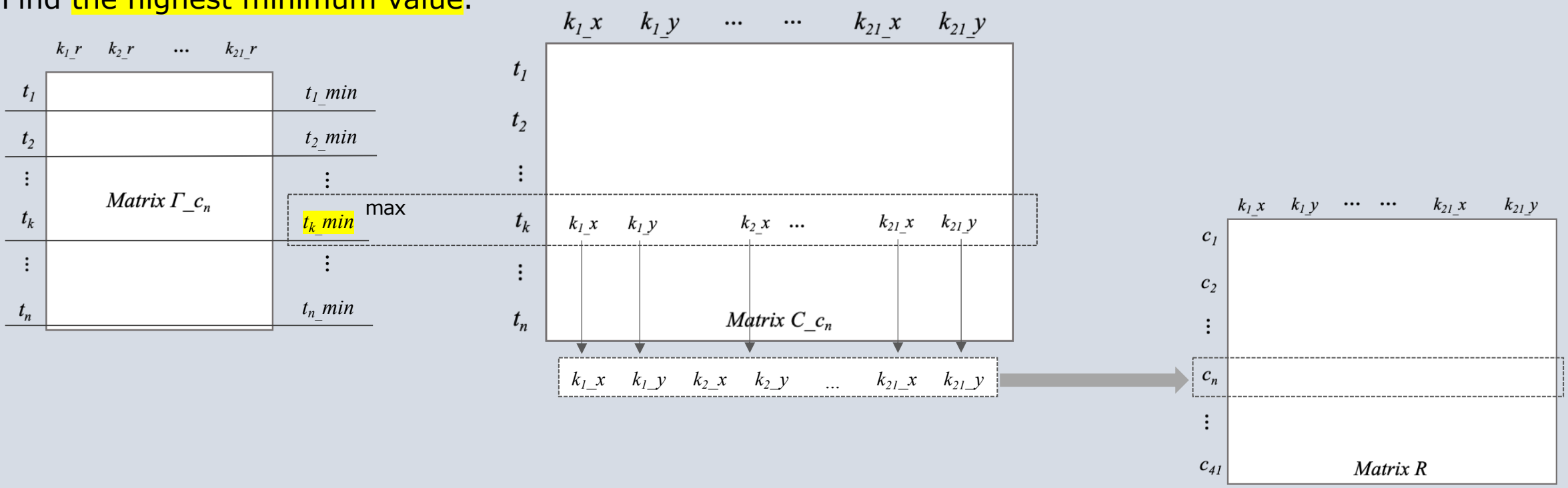
② Creation by **the highest minimum value**

For each frame, extract the minimum value of the reliability.

Find the highest minimum value.

Extract the coordinates of the frame with the highest minimum value.

Store the extracted representative values in the knowledge base



③ Creation by the average of averages

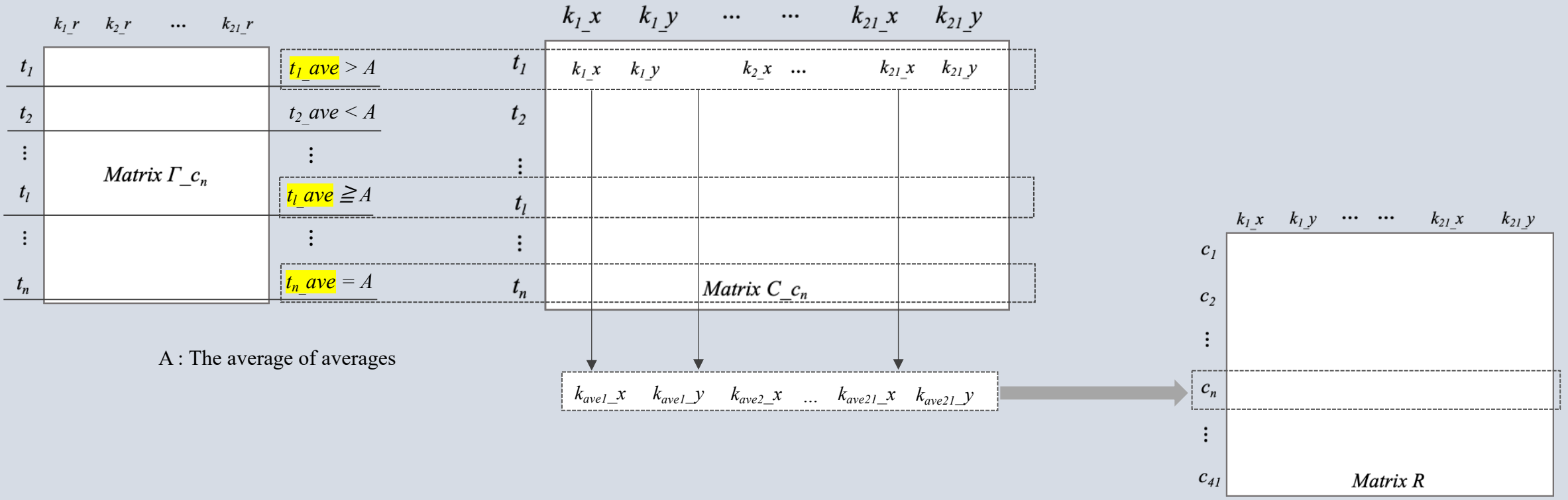
For each frame, calculate the average of the reliability.

Calculate the average value of all averages (= A)

Find all averages higher than A.

Calculate the average for the same coordinates of all extracted frames.

Store the extracted representative values in the knowledge base



A : The average of averages

④ Creation by the sum of the maximum and minimum values

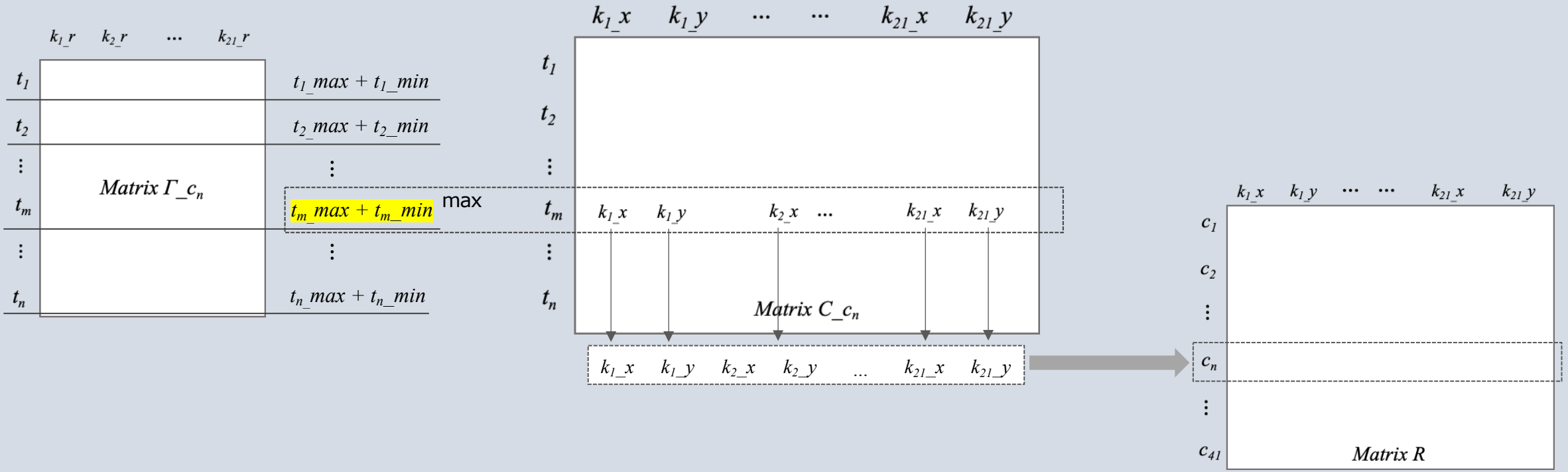
For each frame, extract the maximum and minimum values of the reliability.

Sum the maximum and minimum values.

Find the highest value of the sum.

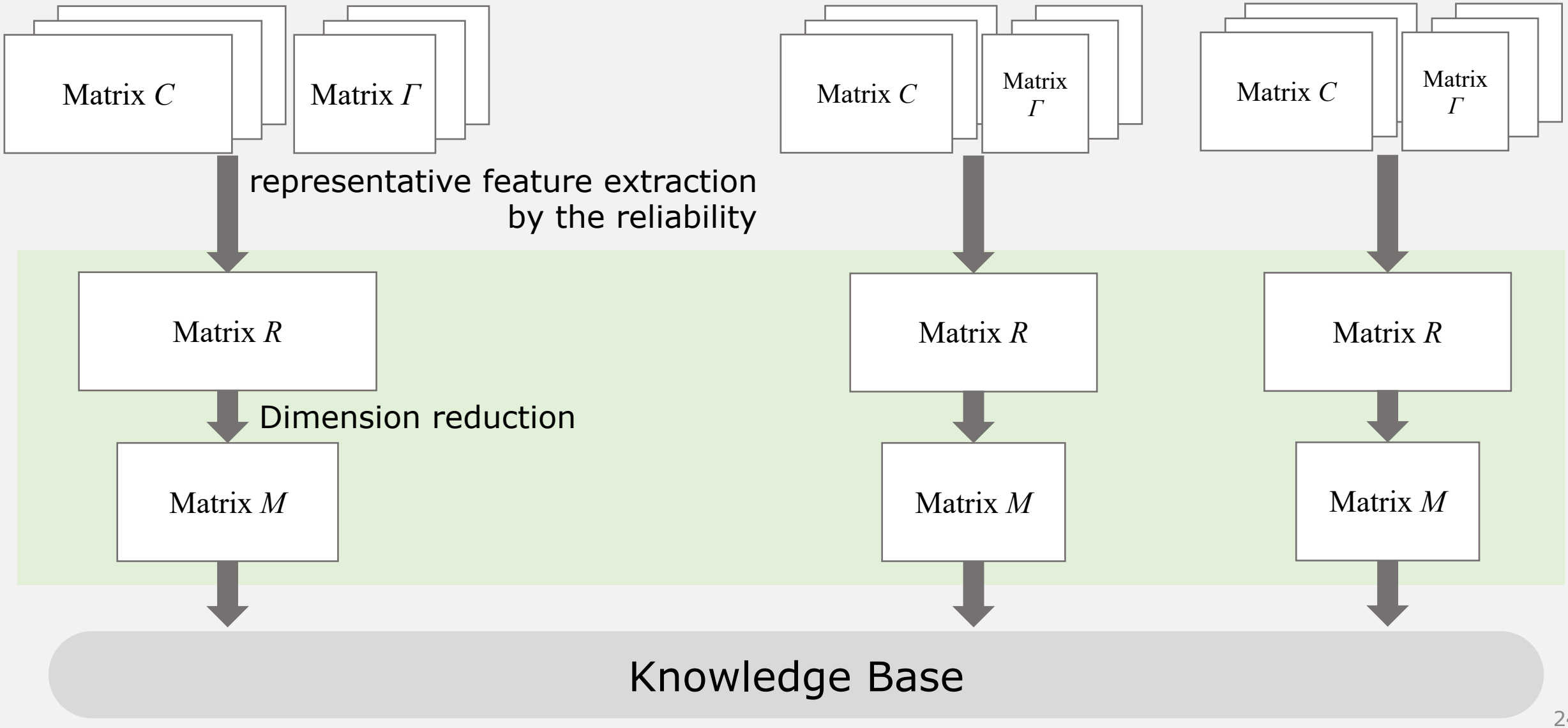
Extract the coordinates of the extracted frame.

Store the extracted representative values in the knowledge base



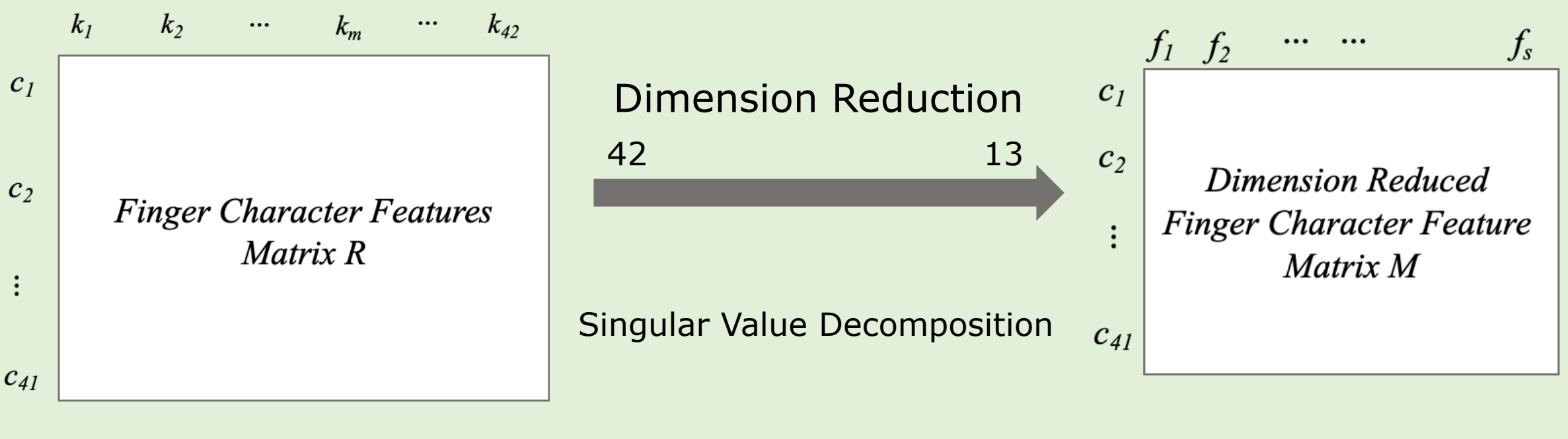
Outline of new knowledge base creation methods by reliability

Proposed method



Dimension reduction

Proposed method



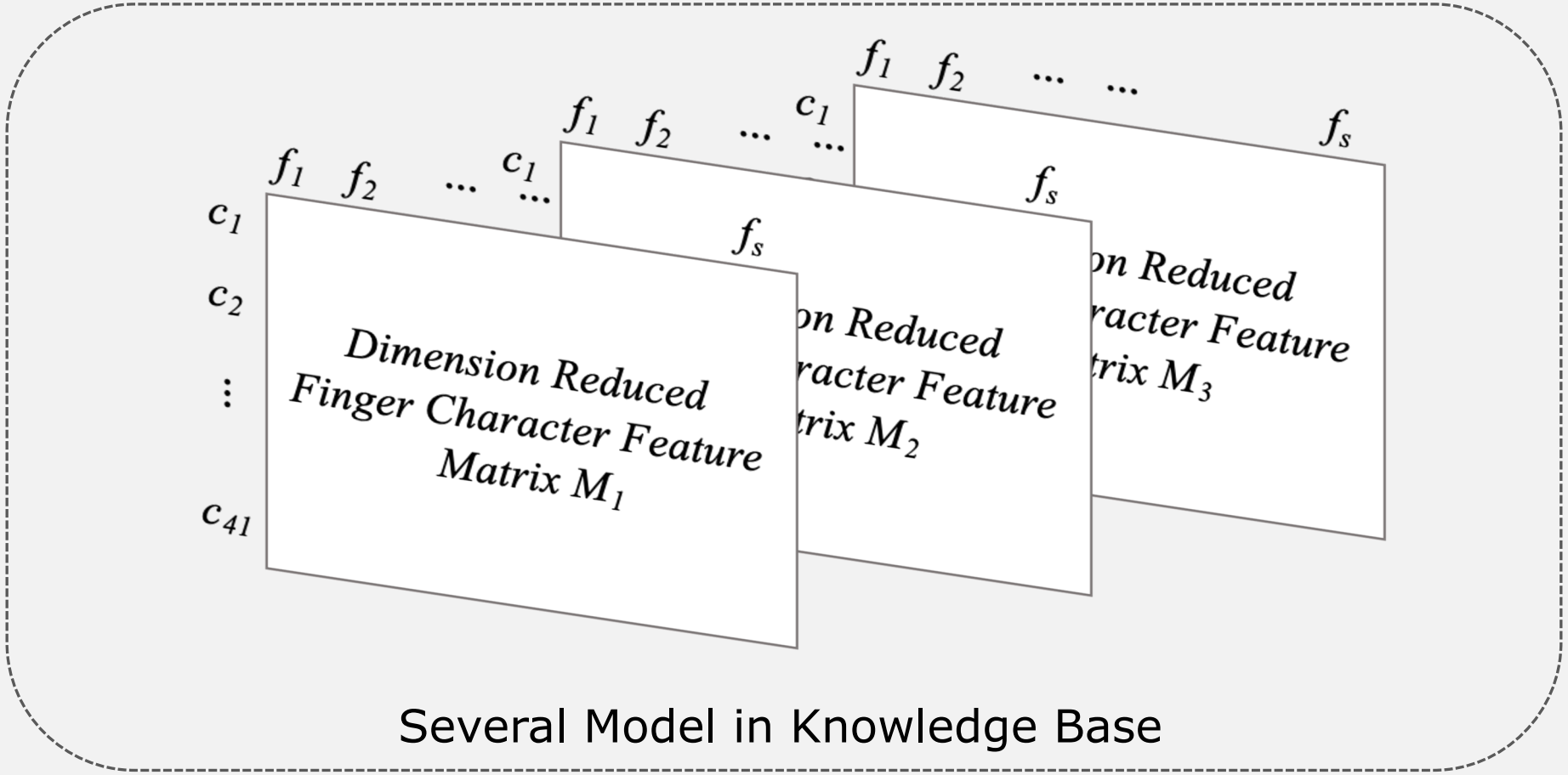
$k_{1\text{-}42}$: X and Y coordinates for each of the 21 finger joint coordinates

$c_{1\text{-}n}$: Each finger character

$f_{1\text{-}s}$: The dimension reduced finger character features

Nitta, T., et al, 2020. Finger Character Recognition in Sign Language Using Finger Feature Knowledge Base for Similarity Measure. *3rd IEEE/ IIAI International Congress on Applied Information Technology*. Tokyo, Japan.

The knowledge base



$c_{1\text{-}n}$: Each finger character

$f_{1\text{-}s}$: The dimension reduced finger character features

Nitta, T., et al, 2020. Finger Character Recognition in Sign Language Using Finger Feature Knowledge Base for Similarity Measure. *3rd IEEE/ IIAI International Congress on Applied Information Technology*. Tokyo, Japan.

Recognition system

Proposed method

A Knowledge Base Creation

Finger Character Video

Annotation

Feature Extraction using OpenPose

Feature Normalization

temporal

temporal

Labeled Temporal Finger Character Features

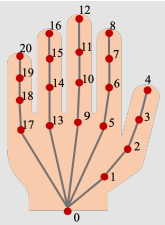
Representative Feature Dimension Reduction

Knowledge Base

Recognition System

Videos and images by normal camera

Feature Extraction using OpenPose



Feature Normalization

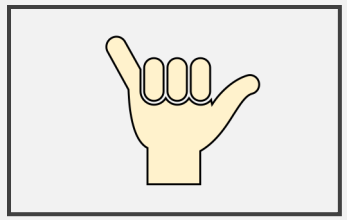
Data Screening

Anomaly Detection
Dimension Reduction

Similarity Measure

Result of Recognition

Similarity measure



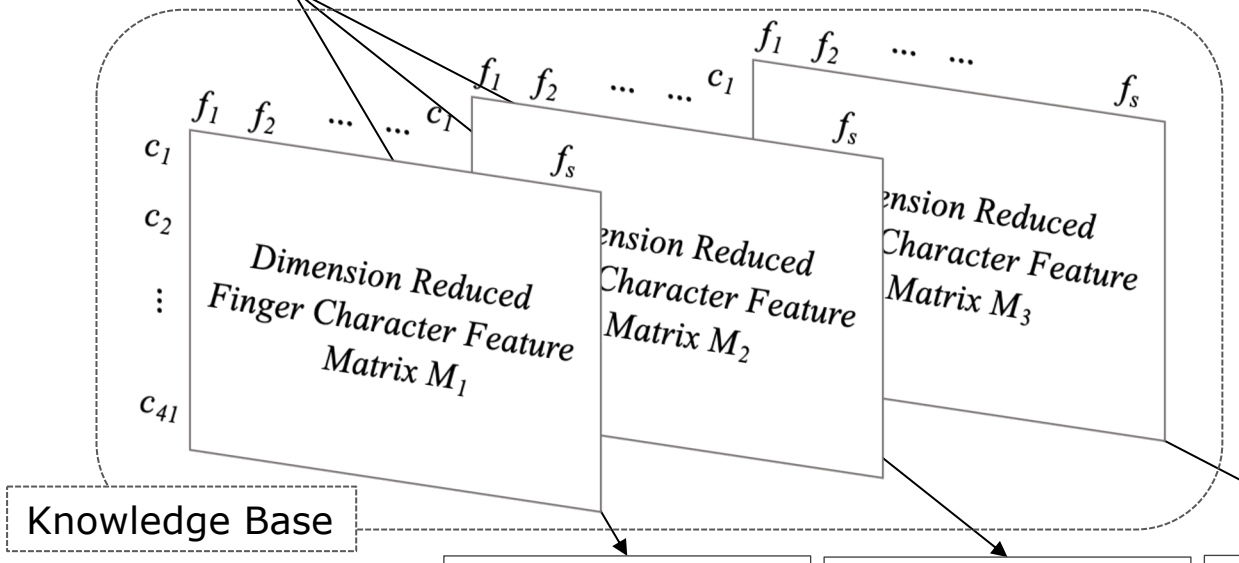
Extraction features

$k_1, k_2, \ldots, k_m, \ldots, k_{42}$

Dimension Reduction

$f_1, f_2, \ldots, f_s$ input data

Similarity Measure



Knowledge Base

"YA" : 99.21 %

"YA" : 94.38 %

"KI" : 97.47 %

Output : "YA", 99.21 %

$k_{1\text{-}42}$: X and Y coordinates for each of the 21 finger joint coordinates

$f_{1\text{-}s}$: The dimension reduced finger character features

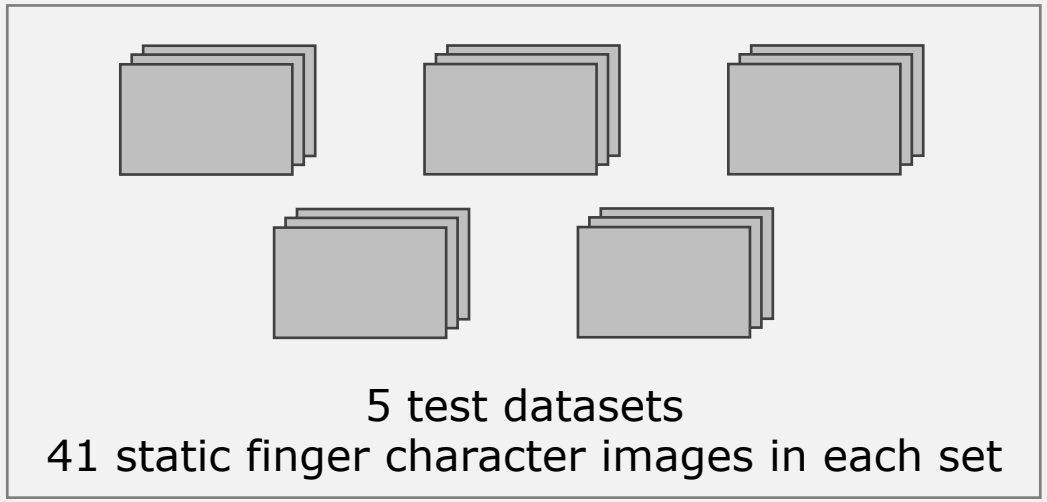
Outline

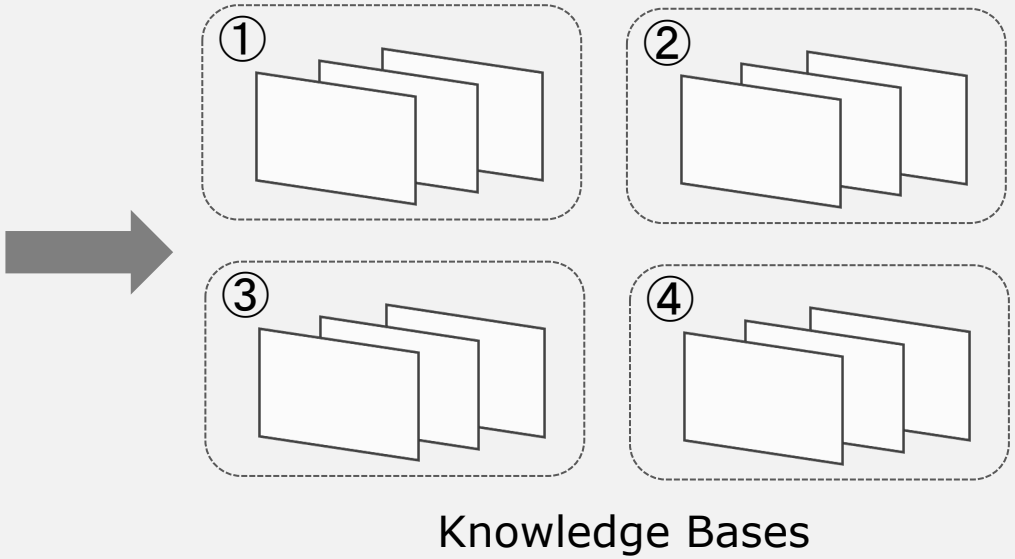
- Background
- Purpose
- Proposed method
- Experiment and Evaluation
- Conclusion

Experiment and evaluation

Experiment

Verify accuracy using test data and knowledge base.





Knowledge Bases

Evaluation

Evaluate and compare with recognition accuracy of previous work.

Experiment result

Perform recognition on each knowledge base and each of the five test data sets.
Calculate the accuracy for each test data set.

Types of Knowledge base	Person-1	Person-2	Person-3	Person-4	Person-5	average
① Highest-Ave	87.50 %	63.41 %	60.98 %	68.75 %	62.50 %	68.63 %
② Highest-Min	82.50 %	63.41 %	60.98 %	65.63 %	62.50 %	67.00 %
③ Average-Ave	85.00 %	68.29 %	58.54 %	71.86 %	60.00 %	68.74 %
④ Sum-of-Max-Min	82.50 %	63.41 %	58.54 %	68.75 %	60.00 %	66.64 %

previous work	Person-1	Person-2	Person-3	Person-4	Person-5	average
accuracy	87.50 %	63.41 %	60.98 %	75.00 %	62.50 %	69.88 %

Person-n : The data consists 41 static finger characters in each set.

Nitta, T., et al, 2020. Finger Character Recognition in Sign Language Using Finger Feature Knowledge Base for Similarity Measure. *3rd IEEE/IIAI International Congress on Applied Information Technology*. Tokyo, Japan.

Discussion

Results of our previous work and the results of the third approach (Average-Ave) that had the highest accuracy in this work.

	Person-1	Person-2	Person-3	Person-4	Person-5	Average
Previous work	87.50 %	63.41 %	60.98 %	75.00 %	62.50 %	69.88 %
③ Average-Ave	85.00 %	68.29 %	58.54 %	71.86 %	60.00 %	68.74 %

About 5%

Positive case

About 3%

Negative case

Example of positive cases

previous work

E

KE

previous work

SHI

MU

③ Average-Ave

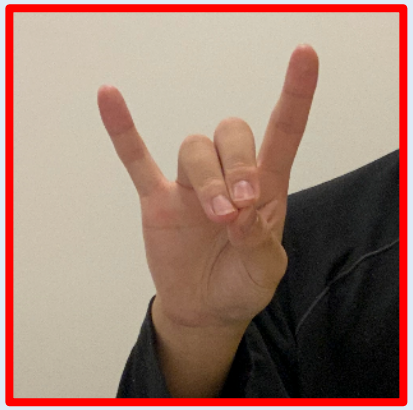
E

③ Average-Ave

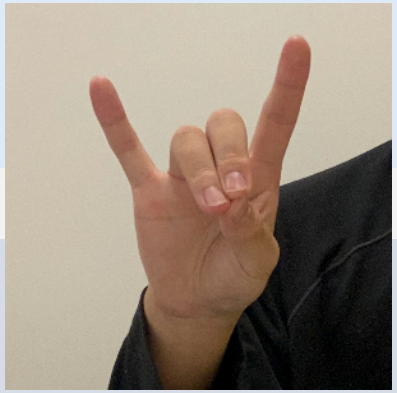
SHI

Example of a negative case

previous work

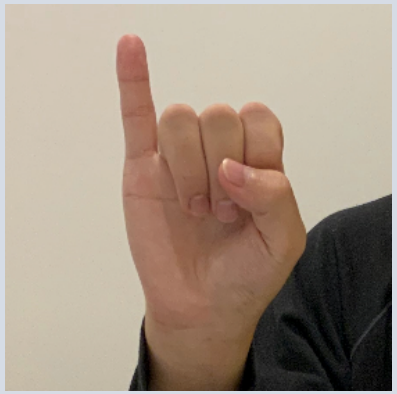


KI



KI

③ Average-Ave



I

Outline

- Background
- Purpose
- Proposed method
- Experiment and Evaluation
- Conclusion

Conclusion

- We proposed a new knowledge base creation method using the reliability of coordinates detected from finger character videos for recognition.
- Our method enables the creation of a knowledge base and metadata with high accuracy from a small amount of training data.
- Our method achieves the same level of accuracy when comparing to our previous work in the finger character recognition using the knowledge base created in our method.

Future work

Ⅰ To improve the accuracy of the recognition

Ⅱ To develop the recognition of dynamic finger characters and sign language that has various motions

Ⅲ To develop an application