

CLAIMS

1. A method, comprising:
 - receiving a video;
 - extracting image data and text data from the video;
 - identifying at least two entities in the image data;
 - assigning at least one entity relation to the at least two entities in the image data;
 - identifying at least two entities in the text data;
 - assigning at least one entity relation to the at least two entities in the text data;
 - generating an image knowledge graph for the at least one entity relation assigned to the at least two entities in the image data;
 - generating a text knowledge graph for the at least one entity relation assigned to the at least two entities in the text data; and
 - generating a weighted knowledge graph based on the image knowledge graph and the text knowledge graph.
2. The method of claim 1, wherein the weighted knowledge graph includes relation weights for the at least one entity relation assigned to the at least two entities in the image data and the at least one entity relation assigned to the at least two entities in the text data.
3. The method of claim 2, further comprising:
 - identifying a top relation in the at least one entity relation assigned to the at least two entities in the image data and the at least one entity relation assigned to the at least two entities in the text data, wherein the top relation is an entity relation having a relation weight greater than a threshold relation weight;
 - selecting frames of the video that correspond to the top relation; and
 - grouping the frames into a video segment.
4. The method of claim 3, further comprising:
 - determining that there are remaining frames of the video that do not include the top relation;
 - determining that the frames in the video segment are nearest to the remaining frames; and
 - grouping the remaining frames with the video segment.
5. The method of claim 1, wherein the video is divided into pictures, wherein each picture includes a set of frames.

6. The method of claim 1, wherein the text data is captions.
7. The method of claim 1, wherein the at least two entities in the image data are identified based on facial recognition.
8. A system comprising means adapted for carrying out all the steps of the method according to any preceding method claim.
9. A computer program comprising instructions for carrying out all the steps of the method according to any preceding method claim, when said computer program is executed on a computer system.