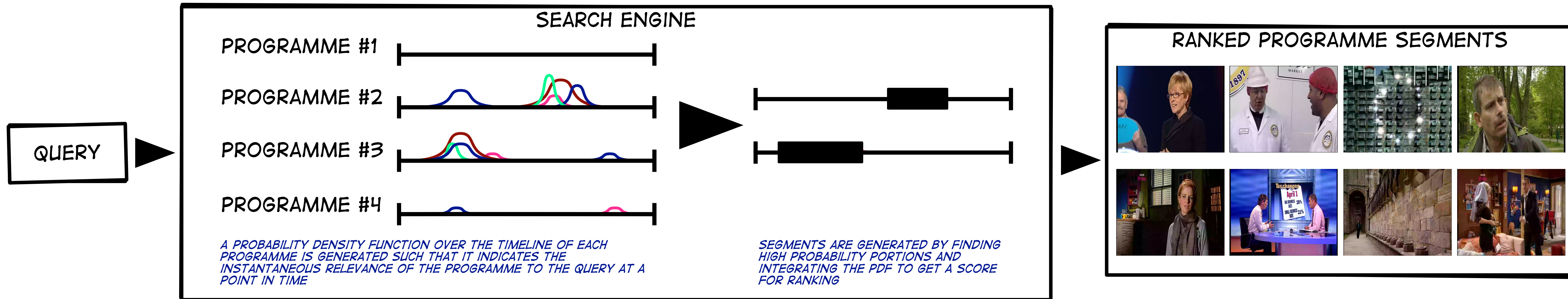


OVERALL APPROACH



SEARCH ENGINE ARCHITECTURE

THE SEARCH ENGINE WAS BUILT IN A MODULAR WAY. DIFFERENT MODULES ARE RUN IN TURN, MODIFYING THE PDF OF EACH PROGRAMME BASED ON THE QUERY.

GENERATING MODULES

GENERATING MODULES INCREASE THE PROBABILITY AT SPECIFIC POINTS ON A TIMELINE BY ADDING GAUSSIAN DISTRIBUTIONS.

TRANSCRIPT SEARCH

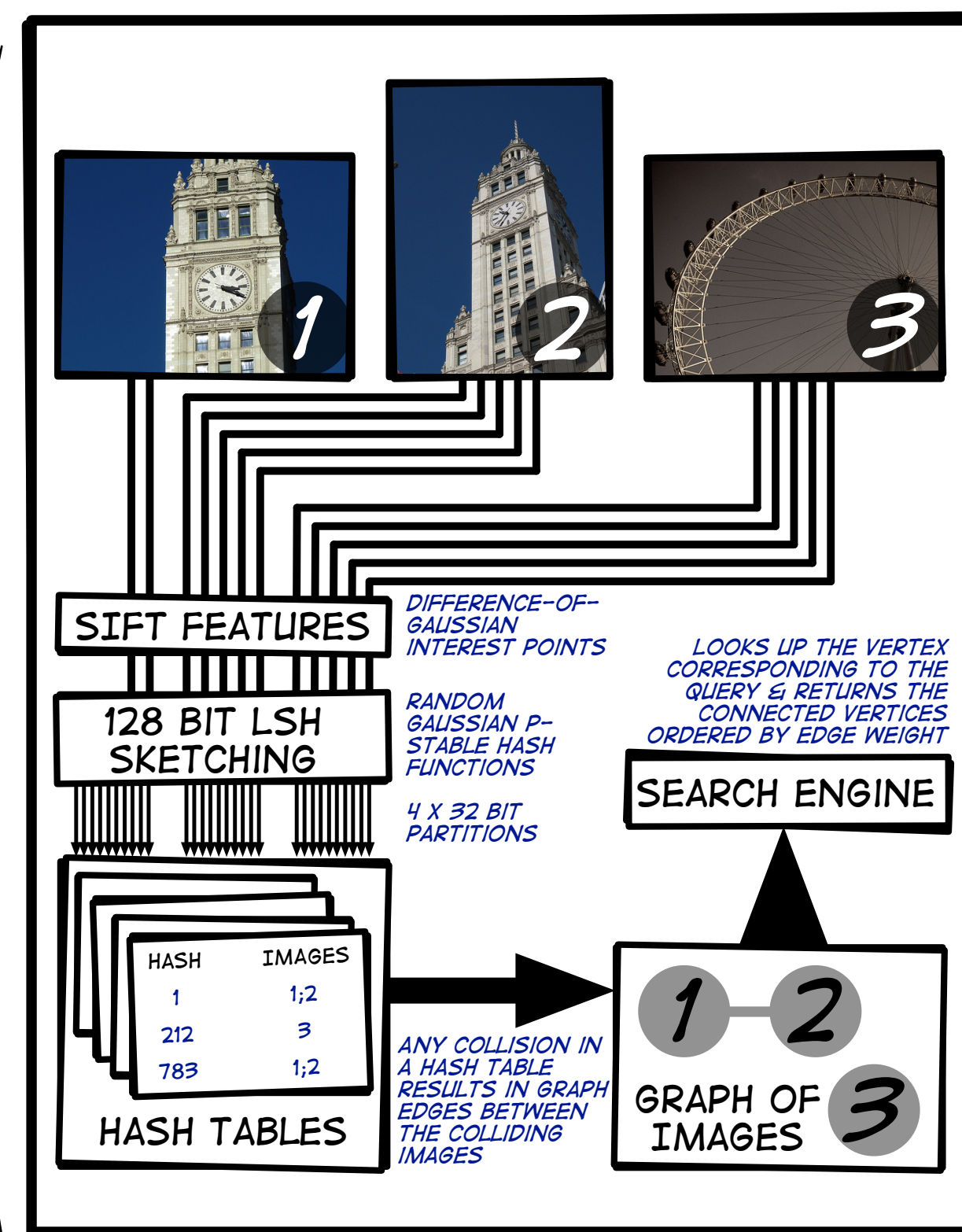
THE TRANSCRIPT IS SEARCHED FOR TERMS IN THE QUERY, AND HITS GENERATE NEW GAUSSIANS AT THE CORRESPONDING POINT ON THE TIMELINE, WITH THE HEIGHT OF THE GAUSSIAN PROPORTIONAL TO THE IDF OF THE TERM.

VISUAL CONCEPT SEARCH

VISUAL CONCEPTS INFERRED FROM THE QUERY ARE MATCHED AGAINST EACH PROGRAMME, CAUSING A GAUSSIAN TO BE PLACED ON THE TIMELINE FOR EVERY HIT.

NEAR DUPLICATE KEYFRAMES

SIMILAR KEYFRAMES ARE MATCHED USING A TECHNIQUE BASED ON LOCALITY SENSITIVE HASHING. SIMILAR KEYFRAMES GENERATE GAUSSIANS AT THE RESPECTIVE TIME WITH HEIGHT PROPORTIONAL TO SIMILARITY.



RESULTS

SEARCH TASK

VISUAL CONCEPTS INFERRED FROM THE QUERY. NEAR-DUPLICATE KEYFRAMES USED FOR QUERY EXPANSION ONCE THE INITIAL QUERY WAS PERFORMED.

RUN TYPE	MRR	MGAP	MASP
SUBTITLES ONLY	0.21	0.10	0.11
TRANSCRIPT +CONCEPTS	0.15	0.07	0.06
TRANSCRIPT +CONCEPTS+LSH	0.12	0.07	0.05

HYPERLINKING TASK

QUERY VISUAL CONCEPTS EXTRACTED FROM THE ANCHOR. ANCHOR KEYFRAMES USED TO FIND NEAR DUPLICATES TARGETS. ANCHOR SUBTITLES USED AS THE QUERY TEXT FOR THE TRANSCRIPT SEARCH MODULE.

RUN TYPE	N.RET	P5	P10	P20	MAP
SUBTITLES ONLY	5393	0.42	0.35	0.22	0.07
TRANSCRIPT +CONCEPTS	7489	0.35	0.30	0.20	0.06
TRANSCRIPT +CONCEPTS+LSH	7488	0.35	0.30	0.21	0.06

WEIGHTING MODULES

WEIGHTING MODULES CHANGE THE OVERALL PDF OF A PROGRAMME BY SCALING ANY ADDED GAUSSIAN FUNCTIONS UP OR DOWN.

SYNOPSIS WEIGHTING

IF TERMS FROM THE TITLE AND/OR SYNOPSIS OF THE PROGRAMME MATCH THE QUERY, THEN THE OVERALL WEIGHT IS INCREASED GIVING THE SEGMENTS FROM THE PROGRAMME A HIGHER PROBABILITY AND RANKING POSITION.

TITLE WEIGHTING

CHANNEL NAME WEIGHTING

IF THE QUERY MENTIONS A CHANNEL (I.E. "BBC THREE") AND THE PROGRAMME WAS NOT BROADCAST ON THAT CHANNEL ITS PDF IS WEIGHTED TO NEAR ZERO, AND THE PROGRAMME WILL NOT APPEAR IN THE RANKED RESULTS.

TITLE **A UNIFIED, MODULAR AND MULTIMODAL APPROACH TO SEARCH AND HYPERLINKING VIDEO**

CLIENT **MEDIAEVAL 2013 SEARCH AND HYPERLINKING OF TELEVISION CONTENT**

POWERED BY **Lucene** **OpenIMAJ**

TEAM **SOTON-WAIS (UNIVERSITY OF SOUTHAMPTON)**

MEMBERS **JOHN PRESTON, JONATHON HARE, SINA SAMANGOUEI, JAMIE DAVIES, NEHA JAIN, DAVID DUPPLAW, PAUL LEWIS**

CONTACT **JONATHON HARE JSH2@ECS.SOTON.AC.UK**

SPONSORS **Southampton UNIVERSITY OF** **arc mem** **Trend Miner**