



MediaEval Benchmark 2014

MediaEval Benchmarking Initiative for Multimedia Evaluation

The "multi" in multimedia: speech, audio, visual content, tags, users, context

Retrieving Diverse Social Images Task

- task overview -

Bogdan Ionescu (UPB, Romania)

Adrian Popescu (CEA LIST, France)

Mihai Lupu (TUW, Austria)

Henning Müller (HES-SO in Sierre, Switzerland)



University Politehnica
of Bucharest



TECHNISCHE
UNIVERSITÄT
WIEN
Vienna University of Technology



October 16-17, Barcelona, Spain

Outline

- The Retrieving Diverse Social Images Task
- Dataset and Evaluation
- Participants
- Results
- Discussion and Perspectives

Diversity Task: Objective & Motivation

Objective: the task addresses the problem of image search **result diversification** in the context of *social photo retrieval*.

Why diversifying search results?

- a method of tackling queries with unclear information needs;
- queries involve many declinations, e.g., sub-topics;
- widens the pool of possible results and increases the system performance;

...

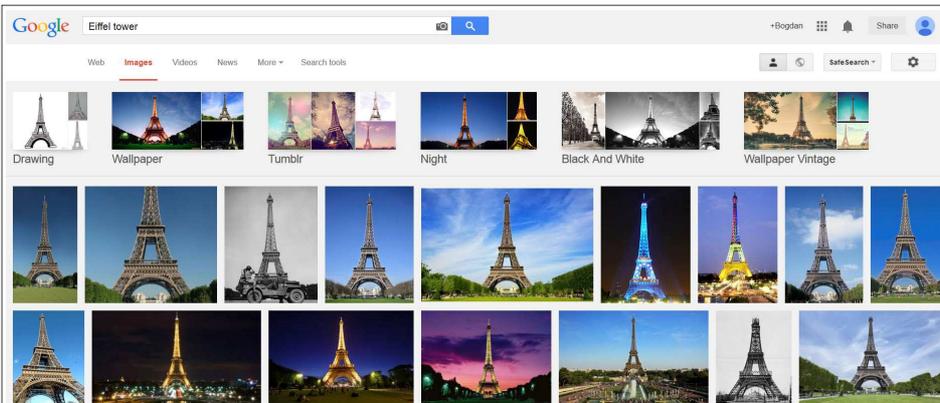
Relevance and Diversity (~antinomic):

too much *diversification* may result in losing relevant items while increasing only the *relevance* will tend to provide near duplicate information.

3

Diversity Task: Objective & Motivation #2

The concept appeared initially for text retrieval but regains its popularity in the context of multimedia retrieval.



[Google Image Search for "Eiffel tower", 12-10-2014]

4

Diversity Task: Use Case

To disambiguate the diversification need, we introduced a very focused use case scenario ...

Use case: we consider a **tourist use case** where a person tries to find more information about a place she is potentially visiting. The person has only a vague idea about the location, knowing the name of the place.

... e.g., looking for **Rialto Bridge** in Italy

5

Diversity Task: Use Case #2

... learn more information from Wikipedia



WIKIPEDIA
The Free Encyclopedia

Rialto Bridge

From Wikipedia, the free encyclopedia

The **Rialto Bridge** (Italian: **Ponte di Rialto**) is one of the four bridges spanning the Grand Canal in Venice, Italy. It is the oldest bridge across the canal, and was the dividing line for the districts of San Marco and San Polo.

Contents	[hide]
1	History
2	See also
3	References
4	External links

History



Two gondoliers pull out with clients on board from a row of gondolas on the Grand Canal near Rialto Bridge.



Detail of the bridge

The first dry crossing of the Grand Canal was a pontoon bridge built in 1181 by Nicolò Barattieri. It was called the Ponte della Moneta, presumably because of the mint that stood near its eastern entrance.^[2]

The development and importance of the Rialto market on the eastern bank increased traffic on the floating bridge, so it was replaced in 1255 by a wooden bridge.^[2] This structure had two inclined ramps meeting at a movable central section, that could be raised to allow the passage of tall ships. The connection with the market eventually led to a change of name for the bridge. During the first half of the 15th century, two rows of shops were built along the sides of the bridge. The rents brought an income to the State Treasury, which helped maintain the bridge.

Maintenance was vital for the timber bridge. It was partly burnt in the revolt led by Bajamonte Tiepolo in 1310. In 1444, it collapsed under the weight of a crowd watching a boat parade and it collapsed again in 1524.

The idea of rebuilding the bridge in stone was first proposed in 1503. Several projects were considered over the following decades. In 1501, the authorities requested proposals for the renewal of the Rialto Bridge, among other things. Plans were offered by famous architects, such as Jacopo Sansovino, Palladio and Vignola, but all involved a Classical approach with several arches, which was judged inappropriate to the situation. Michelangelo also was considered as designer of the bridge.

The present stone bridge, a single span designed by Antonio da Ponte, was finally completed in 1591. It is similar to the wooden bridge it succeeded. Two inclined ramps lead up to a central portico. On either side of the portico, the covered ramps carry rows of shops. The engineering of the bridge was considered so audacious that architect Vincenzo Scamozzi predicted future ruin. The bridge has defied its critics to become one of the architectural icons of Venice.

See also

Miracle of the Relic of the Cross at the Ponte di Rialto (depiction of wooden bridge)

References

- ^[a] Fulton, Charles Carroll (1874). *Europe Viewed Through American Spectacles* (Google books). Philadelphia: J.P. Lippincott & Co. p. 242. Retrieved 2008-09-05. "There being no vehicles or horses in Venice, it is simply for pedestrians."
- ^[a] ^[b] ^[c] ^[d] ^[e] ^[f] ^[g] ^[h] ^[i] ^[j] ^[k] ^[l] ^[m] ^[n] ^[o] ^[p] ^[q] ^[r] ^[s] ^[t] ^[u] ^[v] ^[w] ^[x] ^[y] ^[z] ^[aa] ^[ab] ^[ac] ^[ad] ^[ae] ^[af] ^[ag] ^[ah] ^[ai] ^[aj] ^[ak] ^[al] ^[am] ^[an] ^[ao] ^[ap] ^[aq] ^[ar] ^[as] ^[at] ^[au] ^[av] ^[aw] ^[ax] ^[ay] ^[az] ^[ba] ^[bb] ^[bc] ^[bd] ^[be] ^[bf] ^[bg] ^[bh] ^[bi] ^[bj] ^[bk] ^[bl] ^[bm] ^[bn] ^[bo] ^[bp] ^[bq] ^[br] ^[bs] ^[bt] ^[bu] ^[bv] ^[bv] ^[bw] ^[bx] ^[by] ^[bz] ^[ca] ^[cb] ^[cc] ^[cd] ^[ce] ^[cf] ^[cf] ^[cg] ^[ch] ^[ci] ^[cj] ^[ck] ^[cl] ^[cm] ^[cn] ^[co] ^[cp] ^[cq] ^[cr] ^[cs] ^[ct] ^[ct] ^[cu] ^[cv] ^[cv] ^[cw] ^[cx] ^[cy] ^[cy] ^[cz] ^[ca] ^[ca] ^[cb] ^[cb] ^[cc] ^[cc] ^[cd] ^[cd] ^[ce] ^[ce] ^[cf] ^[cf] ^[cf] ^[cg] ^[cg] ^[ch] ^[ch] ^[ch] ^[ci] ^[ci] ^[ci] ^[cj] ^[cj] ^[cj] ^[ck] ^[ck] ^[ck] ^[cl] ^[cl] ^[cl] ^[cm] ^[cm] ^[cm] ^[cn] ^[cn] ^[cn] ^[co] ^[co] ^[co] ^[cp] ^[cp] ^[cp] ^[cq] ^[cq] ^[cq] ^[cr] ^[cr] ^[cr] ^[cs] ^[cs] ^[cs] ^[ct] ^[ct] ^[ct] ^[cu] ^[cu] ^[cu] ^[cv] ^[cv] ^[cv] ^[cw] ^[cw] ^[cw] ^[cx] ^[cx] ^[cx] ^[cy] ^[cy] ^[cy] ^[cz] ^[cz] ^[cz] ^[da] ^[da] ^[da] ^[db] ^[db] ^[db] ^[dc] ^[dc] ^[dc] ^[dd] ^[dd] ^[dd] ^[de] ^[de] ^[de] ^[df] ^[df] ^[df] ^[dg] ^[dg] ^[dg] ^[dh] ^[dh] ^[dh] ^[di] ^[di] ^[di] ^[dj] ^[dj] ^[dj] ^[dk] ^[dk] ^[dk] ^[dl] ^[dl] ^[dl] ^[dm] ^[dm] ^[dm] ^[dn] ^[dn] ^[dn] ^[do] ^[do] ^[do] ^[dp] ^[dp] ^[dp] ^[dq] ^[dq] ^[dq] ^[dr] ^[dr] ^[dr] ^[ds] ^[ds] ^[ds] ^[dt] ^[dt] ^[dt] ^[du] ^[du] ^[du] ^[dv] ^[dv] ^[dv] ^[dw] ^[dw] ^[dw] ^[dx] ^[dx] ^[dx] ^[dy] ^[dy] ^[dy] ^[dz] ^[dz] ^[dz] ^[ea] ^[ea] ^[ea] ^[eb] ^[eb] ^[eb] ^[ec] ^[ec] ^[ec] ^[ed] ^[ed] ^[ed] ^[ee] ^[ee] ^[ee] ^[ef] ^[ef] ^[ef] ^[eg] ^[eg] ^[eg] ^[eh] ^[eh] ^[eh] ^[ei] ^[ei] ^[ei] ^[ej] ^[ej] ^[ej] ^[ek] ^[ek] ^[ek] ^[el] ^[el] ^[el] ^[em] ^[em] ^[em] ^[en] ^[en] ^[en] ^[eo] ^[eo] ^[eo] ^[ep] ^[ep] ^[ep] ^[eq] ^[eq] ^[eq] ^[er] ^[er] ^[er] ^[es] ^[es] ^[es] ^[et] ^[et] ^[et] ^[eu] ^[eu] ^[eu] ^[ev] ^[ev] ^[ev] ^[ew] ^[ew] ^[ew] ^[ex] ^[ex] ^[ex] ^[ey] ^[ey] ^[ey] ^[ez] ^[ez] ^[ez] ^[fa] ^[fa] ^[fa] ^[fb] ^[fb] ^[fb] ^[fc] ^[fc] ^[fc] ^[fd] ^[fd] ^[fd] ^[fe] ^[fe] ^[fe] ^[ff] ^[ff] ^[ff] ^[fg] ^[fg] ^[fg] ^[fh] ^[fh] ^[fh] ^[fi] ^[fi] ^[fi] ^[fj] ^[fj] ^[fj] ^[fk] ^[fk] ^[fk] ^[fl] ^[fl] ^[fl] ^[fm] ^[fm] ^[fm] ^[fn] ^[fn] ^[fn] ^[fo] ^[fo] ^[fo] ^[fp] ^[fp] ^[fp] ^[fq] ^[fq] ^[fq] ^[fr] ^[fr] ^[fr] ^[fs] ^[fs] ^[fs] ^[ft] ^[ft] ^[ft] ^[fu] ^[fu] ^[fu] ^[fv] ^[fv] ^[fv] ^[fw] ^[fw] ^[fw] ^[fx] ^[fx] ^[fx] ^[fy] ^[fy] ^[fy] ^[fz] ^[fz] ^[fz] ^[ga] ^[ga] ^[ga] ^[gb] ^[gb] ^[gb] ^[gc] ^[gc] ^[gc] ^[gd] ^[gd] ^[gd] ^[ge] ^[ge] ^[ge] ^[gf] ^[gf] ^[gf] ^[gg] ^[gg] ^[gg] ^[gh] ^[gh] ^[gh] ^[gi] ^[gi] ^[gi] ^[gj] ^[gj] ^[gj] ^[gk] ^[gk] ^[gk] ^[gl] ^[gl] ^[gl] ^[gm] ^[gm] ^[gm] ^[gn] ^[gn] ^[gn] ^[go] ^[go] ^[go] ^[gp] ^[gp] ^[gp] ^[gq] ^[gq] ^[gq] ^[gr] ^[gr] ^[gr] ^[gs] ^[gs] ^[gs] ^[gt] ^[gt] ^[gt] ^[gu] ^[gu] ^[gu] ^[gv] ^[gv] ^[gv] ^[gw] ^[gw] ^[gw] ^[gx] ^[gx] ^[gx] ^[gy] ^[gy] ^[gy] ^[gz] ^[gz] ^[gz] ^[ha] ^[ha] ^[ha] ^[hb] ^[hb] ^[hb] ^[hc] ^[hc] ^[hc] ^[hd] ^[hd] ^[hd] ^[he] ^[he] ^[he] ^[hf] ^[hf] ^[hf] ^[hg] ^[hg] ^[hg] ^[hh] ^[hh] ^[hh] ^[hi] ^[hi] ^[hi] ^[hj] ^[hj] ^[hj] ^[hk] ^[hk] ^[hk] ^[hl] ^[hl] ^[hl] ^[hm] ^[hm] ^[hm] ^[hn] ^[hn] ^[hn] ^[ho] ^[ho] ^[ho] ^[hp] ^[hp] ^[hp] ^[hq] ^[hq] ^[hq] ^[hr] ^[hr] ^[hr] ^[hs] ^[hs] ^[hs] ^[ht] ^[ht] ^[ht] ^[hu] ^[hu] ^[hu] ^[hv] ^[hv] ^[hv] ^[hw] ^[hw] ^[hw] ^[hx] ^[hx] ^[hx] ^[hy] ^[hy] ^[hy] ^[hz] ^[hz] ^[hz] ^[ia] ^[ia] ^[ia] ^[ib] ^[ib] ^[ib] ^[ic] ^[ic] ^[ic] ^[id] ^[id] ^[id] ^[ie] ^[ie] ^[ie] ^[if] ^[if] ^[if] ^[ig] ^[ig] ^[ig] ^[ih] ^[ih] ^[ih] ^[ii] ^[ii] ^[ii] ^[ij] ^[ij] ^[ij] ^[ik] ^[ik] ^[ik] ^[il] ^[il] ^[il] ^[im] ^[im] ^[im] ^[in] ^[in] ^[in] ^[io] ^[io] ^[io] ^[ip] ^[ip] ^[ip] ^[iq] ^[iq] ^[iq] ^[ir] ^[ir] ^[ir] ^[is] ^[is] ^[is] ^[it] ^[it] ^[it] ^[iu] ^[iu] ^[iu] ^[iv] ^[iv] ^[iv] ^[iw] ^[iw] ^[iw] ^[ix] ^[ix] ^[ix] ^[iy] ^[iy] ^[iy] ^[iz] ^[iz] ^[iz] ^[ja] ^[ja] ^[ja] ^[jb] ^[jb] ^[jb] ^[jc] ^[jc] ^[jc] ^[jd] ^[jd] ^[jd] ^[je] ^[je] ^[je] ^[jf] ^[jf] ^[jf] ^[jg] ^[jg] ^[jg] ^[jh] ^[jh] ^[jh] ^[ji] ^[ji] ^[ji] ^[jj] ^[jj] ^[jj] ^[jk] ^[jk] ^[jk] ^[jl] ^[jl] ^[jl] ^[jm] ^[jm] ^[jm] ^[jn] ^[jn] ^[jn] ^[jo] ^[jo] ^[jo] ^[jp] ^[jp] ^[jp] ^[jq] ^[jq] ^[jq] ^[jr] ^[jr] ^[jr] ^[js] ^[js] ^[js] ^[jt] ^[jt] ^[jt] ^[ju] ^[ju] ^[ju] ^[jv] ^[jv] ^[jv] ^[jw] ^[jw] ^[jw] ^[jx] ^[jx] ^[jx] ^[jy] ^[jy] ^[jy] ^[jz] ^[jz] ^[jz] ^[ka] ^[ka] ^[ka] ^[kb] ^[kb] ^[kb] ^[kc] ^[kc] ^[kc] ^[kd] ^[kd] ^[kd] ^[ke] ^[ke] ^[ke] ^[kf] ^[kf] ^[kf] ^[kg] ^[kg] ^[kg] ^[kh] ^[kh] ^[kh] ^[ki] ^[ki] ^[ki] ^[kj] ^[kj] ^[kj] ^[kk] ^[kk] ^[kk] ^[kl] ^[kl] ^[kl] ^[km] ^[km] ^[km] ^[kn] ^[kn] ^[kn] ^[ko] ^[ko] ^[ko] ^[kp] ^[kp] ^[kp] ^[kq] ^[kq] ^[kq] ^[kr] ^[kr] ^[kr] ^[ks] ^[ks] ^[ks] ^[kt] ^[kt] ^[kt] ^[ku] ^[ku] ^[ku] ^[kv] ^[kv] ^[kv] ^[kw] ^[kw] ^[kw] ^[kx] ^[kx] ^[kx] ^[ky] ^[ky] ^[ky] ^[kz] ^[kz] ^[kz] ^[la] ^[la] ^[la] ^[lb] ^[lb] ^[lb] ^[lc] ^[lc] ^[lc] ^[ld] ^[ld] ^[ld] ^[le] ^[le] ^[le] ^[lf] ^[lf] ^[lf] ^[lg] ^[lg] ^[lg] ^[lh] ^[lh] ^[lh] ^[li] ^[li] ^[li] ^[lj] ^[lj] ^[lj] ^[lk] ^[lk] ^[lk] ^[ll] ^[ll] ^[ll] ^[lm] ^[lm] ^[lm] ^[ln] ^[ln] ^[ln] ^[lo] ^[lo] ^[lo] ^[lp] ^[lp] ^[lp] ^[lq] ^[lq] ^[lq] ^[lr] ^[lr] ^[lr] ^[ls] ^[ls] ^[ls] ^[lt] ^[lt] ^[lt] ^[lu] ^[lu] ^[lu] ^[lv] ^[lv] ^[lv] ^[lw] ^[lw] ^[lw] ^[lx] ^[lx] ^[lx] ^[ly] ^[ly] ^[ly] ^[lz] ^[lz] ^[lz] ^[ma] ^[ma] ^[ma] ^[mb] ^[mb] ^[mb] ^[mc] ^[mc] ^[mc] ^[md] ^[md] ^[md] ^[me] ^[me] ^[me] ^[mf] ^[mf] ^[mf] ^[mg] ^[mg] ^[mg] ^[mh] ^[mh] ^[mh] ^[mi] ^[mi] ^[mi] ^[mj] ^[mj] ^[mj] ^[mk] ^[mk] ^[mk] ^[ml] ^[ml] ^[ml] ^[mm] ^[mm] ^[mm] ^[mn] ^[mn] ^[mn] ^[mo] ^[mo] ^[mo] ^[mp] ^[mp] ^[mp] ^[mq] ^[mq] ^[mq] ^[mr] ^[mr] ^[mr] ^[ms] ^[ms] ^[ms] ^[mt] ^[mt] ^[mt] ^[mu] ^[mu] ^[mu] ^[mv] ^[mv] ^[mv] ^[mw] ^[mw] ^[mw] ^[mx] ^[mx] ^[mx] ^[my] ^[my] ^[my] ^[mz] ^[mz] ^[mz] ^[na] ^[na] ^[na] ^[nb] ^[nb] ^[nb] ^[nc] ^[nc] ^[nc] ^[nd] ^[nd] ^[nd] ^[ne] ^[ne] ^[ne] ^[nf] ^[nf] ^[nf] ^[ng] ^[ng] ^[ng] ^[nh] ^[nh] ^[nh] ^[ni] ^[ni] ^[ni] ^[nj] ^[nj] ^[nj] ^[nk] ^[nk] ^[nk] ^[nl] ^[nl] ^[nl] ^[nm] ^[nm] ^[nm] ^[nn] ^[nn] ^[nn] ^[no] ^[no] ^[no] ^[np] ^[np] ^[np] ^[nq] ^[nq] ^[nq] ^[nr] ^[nr] ^[nr] ^[ns] ^[ns] ^[ns] ^[nt] ^[nt] ^[nt] ^[nu] ^[nu] ^[nu] ^[nv] ^[nv] ^[nv] ^[nw] ^[nw] ^[nw] ^[nx] ^[nx] ^[nx] ^[ny] ^[ny] ^[ny] ^[nz] ^[nz] ^[nz] ^[oa] ^[oa] ^[oa] ^[ob] ^[ob] ^[ob] ^[oc] ^[oc] ^[oc] ^[od] ^[od] ^[od] ^[oe] ^[oe] ^[oe] ^[of] ^[of] ^[of] ^[og] ^[og] ^[og] ^[oh] ^[oh] ^[oh] ^[oi] ^[oi] ^[oi] ^[oj] ^[oj] ^[oj] ^[ok] ^[ok] ^[ok] ^[ol] ^[ol] ^[ol] ^[om] ^[om] ^[om] ^[on] ^[on] ^[on] ^[oo] ^[oo] ^[oo] ^[op] ^[op] ^[op] ^[oq] ^[oq] ^[oq] ^[or] ^[or] ^[or] ^[os] ^[os] ^[os] ^[ot] ^[ot] ^[ot] ^[ou] ^[ou] ^[ou] ^[ov] ^[ov] ^[ov] ^[ow] ^[ow] ^[ow] ^[ox] ^[ox] ^[ox] ^[oy] ^[oy] ^[oy] ^[oz] ^[oz] ^[oz] ^[pa] ^[pa] ^[pa] ^[pb] ^[pb] ^[pb] ^[pc] ^[pc] ^[pc] ^[pd] ^[pd] ^[pd] ^[pe] ^[pe] ^[pe] ^[pf] ^[pf] ^[pf] ^[pg] ^[pg] ^[pg] ^[ph] ^[ph] ^[ph] ^[pi] ^[pi] ^[pi] ^[pj] ^[pj] ^[pj] ^[pk] ^[pk] ^[pk] ^[pl] ^[pl] ^[pl] ^[pm] ^[pm] ^[pm] ^[pn] ^[pn] ^[pn] ^[po] ^[po] ^[po] ^[pp] ^[pp] ^[pp] ^[pq] ^[pq] ^[pq] ^[pr] ^[pr] ^[pr] ^[ps] ^[ps] ^[ps] ^[pt] ^[pt] ^[pt] ^[pu] ^[pu] ^[pu] ^[pv] ^[pv] ^[pv] ^[pw] ^[pw] ^[pw] ^[px] ^[px] ^[px] ^[py] ^[py] ^[py] ^[pz] ^[pz] ^[pz] ^[qa] ^[qa] ^[qa] ^[qb] ^[qb] ^[qb] ^[qc] ^[qc] ^[qc] ^[qd] ^[qd] ^[qd] ^[qe] ^[qe] ^[qe] ^[qf] ^[qf] ^[qf] ^[qg] ^[qg] ^[qg] ^[qh] ^[qh] ^[qh] ^[qi] ^[qi] ^[qi] ^[qj] ^[qj] ^[qj] ^[qk] ^[qk] ^[qk] ^[ql] ^[ql] ^[ql] ^[qm] ^[qm] ^[qm] ^[qn] ^[qn] ^[qn] ^[qo] ^[qo] ^[qo] ^[qp] ^[qp] ^[qp] ^[qq] ^[qq] ^[qq] ^[qr] ^[qr] ^[qr] ^[qs] ^[qs] ^[qs] ^[qt] ^[qt] ^[qt] ^[qu] ^[qu] ^[qu] ^[qv] ^[qv] ^[qv] ^[qw] ^[qw] ^[qw] ^[qx] ^[qx] ^[qx] ^[qy] ^[qy] ^[qy] ^[qz] ^[qz] ^[qz] ^[ra] ^[ra] ^[ra] ^[rb] ^[rb] ^[rb] ^[rc] ^[rc] ^[rc] ^[rd] ^[rd] ^[rd] ^[re] ^[re] ^[re] ^[rf] ^[rf] ^[rf] ^[rg] ^[rg] ^[rg] ^[rh] ^[rh] ^[rh] ^[ri] ^[ri] ^[ri] ^[rj] ^[rj] ^[rj] ^[rk] ^[rk] ^[rk] ^[rl] ^[rl] ^[rl] ^[rm] ^[rm] ^[rm] ^[rn] ^[rn] ^[rn] ^[ro] ^[ro] ^[ro] ^[rp] ^[rp] ^[rp] ^[rq] ^[rq] ^[rq] ^[rr] ^[rr] ^[rr] ^[rs] ^[rs] ^[rs] ^[rt] ^[rt] ^[rt] ^[ru] ^[ru] ^[ru] ^[rv] ^[rv] ^[rv] ^[rw] ^[rw] ^[rw] ^[rx] ^[rx] ^[rx] ^[ry] ^[ry] ^[ry] ^[rz] ^[rz] ^[rz] ^[sa] ^[sa] ^[sa] ^[sb] ^[sb] ^[sb] ^[sc] ^[sc] ^[sc] ^[sd] ^[sd] ^[sd] ^[se] ^[se] ^[se] ^[sf] ^[sf] ^[sf] ^[sg] ^[sg] ^[sg] ^[sh] ^[sh] ^[sh] ^[si] ^[si] ^[si] ^[sj] ^[sj] ^[sj] ^[sk] ^[sk] ^[sk] ^[sl] ^[sl] ^[sl] ^[sm] ^[sm] ^[sm] ^[sn] ^[sn] ^[sn] ^[so] ^[so] ^[so] ^[sp] ^[sp] ^[sp] ^[sq] ^[sq] ^[sq] ^[sr] ^[sr] ^[sr] ^[ss] ^[ss] ^[ss] ^[st] ^[st] ^[st] ^[su] ^[su] ^[su] ^[sv] ^[sv] ^[sv] ^[sw] ^[sw] ^[sw] ^[sx] ^[sx] ^[sx] ^[sy] ^[sy] ^[sy] ^[sz] ^[sz] ^[sz] ^[ta] ^[ta] ^[ta] ^[tb] ^[tb] ^[tb] ^[tc] ^[tc] ^[tc] ^[td] ^[td] ^[td] ^[te] ^[te] ^[te] ^[tf] ^[tf] ^[tf] ^[tg] ^[tg] ^[tg] ^[th] ^[th] ^[th] ^[ti] ^[ti] ^[ti] ^[tj] ^[tj] ^[tj] ^[tk] ^[tk] ^[tk] ^[tl] ^[tl] ^[tl] ^[tm] ^[tm] ^[tm] ^[tn] ^[tn] ^[tn] ^[to] ^[to] ^[to] ^[tp] ^[tp] ^[tp] ^[tq] ^[tq] ^[tq] ^[tr] ^[tr] ^[tr] ^[ts] ^[ts] ^[ts] ^[tt] ^[tt] ^[tt] ^[tu] ^[tu] ^[tu] ^[tv] ^[tv] ^[tv] [[]

Diversity Task: Use Case #3

... how to get some more accurate photos ?



query using text "Rialto Bridge" ...

... browse the results

7

Diversity Task: Use Case #4



page 1

8

Diversity Task: Use Case #5



page n

9

Diversity Task: Use Case #6

... too many results to process,

inaccurate, e.g., people in focus, other views or places



meaningless objects



redundant results, e.g., duplicates, similar views ...



10

Diversity Task: Use Case #7



page 1

11

Diversity Task: Use Case #8



page n

12

Diversity Task: Definition

Participants receive a *ranked list* of photos with locations retrieved from Flickr using its default “relevance” algorithm.

Goal of the task: *refine* the results by providing a *ranked* list of *up to 50 photos (summary)* that are considered to be both *relevant* and *diverse* representations of the query.

relevant*: common photo representation of the location, e.g., different views at different times of the day/year and under different weather conditions, inside views, close-ups, drawings, sketches, creative views, which contain partially or entirely the target location.

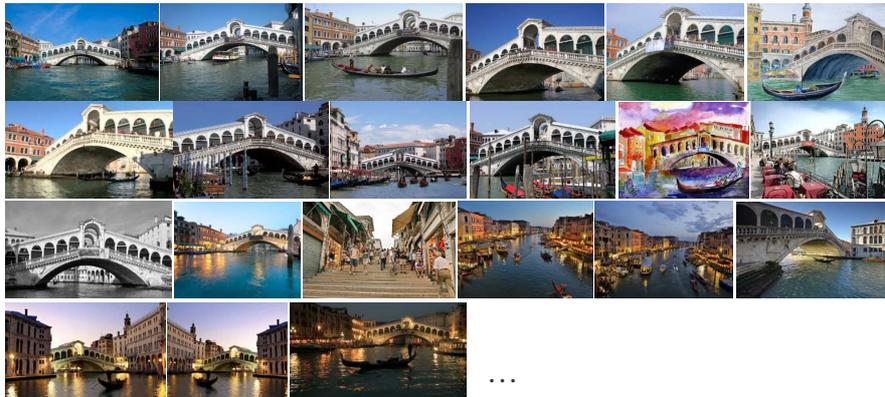
diverse*: depicting different visual characteristics of the location, with a certain degree of complementarity, i.e., most of the perceived visual information is different from one photo to another.

*we thank the task survey respondents for their precious feedback on these definitions.

13

Diversity Task: Target

going from this ...



14

Diversity Task: Target

... to something like this:



15

Dataset: General Information

The dataset consists of **300 landmark locations** (natural or man-made, e.g., sites, museums, monuments, buildings, roads, bridges) unevenly spread over 35 countries around the world:



16

Dataset: Resources

Location information consists of:

- the location name & GPS coordinates;
- a link to its Wikipedia web page;
- up to **5 representative photos** from Wikipedia;
- a ranked set of Creative Commons photos retrieved from Flickr (up to **300 photos per location**);
- metadata from Flickr (e.g., tags, description, views, #comments, date-time photo was taken, username, **userid**, etc);
- some general purpose visual and text content descriptors;
- **an automatic prediction of user annotation credibility**;
- relevance and diversity ground truth (**up to 25 classes**).

Retrieval method (we use Flickr API):

- use of the location name as query.

[2014: more focus on social aspects]

* the differences compared to 2013 data are depicted in bold.

17

Dataset: User Credibility

Idea: give an automatic estimation of the quality of tag-image content relationships;

~ indication about which users are most likely to share relevant images in Flickr (according to the underlying task scenario).

- **visualScore:** for each Flickr tag which is identical to an ImageNet concept, a classification score is predicted and the visualScore of a user is obtained by averaging individual tag scores;

- **faceProportion:** the percentage of images with faces out of the total of images tested for each user;

- **uploadFrequency:** average time between two consecutive uploads in Flickr;

...

18

Dataset: Statistics

Some basic statistics:

- **devset** (intended for designing and validating the methods)

<i>#locations</i>	<i>#images</i>	<i>min-average-max img. per location</i>
30	8,923	285 - 297 - 300

- **testset** (intended for final benchmark)

<i>#locations</i>	<i>#images</i>	<i>min-average-max img. per location</i>
123	36,452	277 - 296 - 300

⇒ total number of provided images: 45,375.

- **credibilityset** (intended for training/designing credibility desc.)

<i>#locations</i>	<i>#images*</i>	<i>#users</i>	<i>average img. per user</i>
300	3,651,303	685	5,330

* images are provided via Flickr URLs.

19

Dataset: Ground Truth

Relevance and diversity annotations were carried out by **expert annotators***:

- **devset**: relevance (3 annotations), diversity (1 annotation issued from 2 experts + 1 final master revision);
- **testset**: relevance (3 annotations issued from 11 expert annotators), diversity (1 annotation from 3 expert annotators + 1 final master revision);
- **credibilityset**: only relevance for 50,157 photos (3 annotations issued from 9 experts);
- lenient majority voting for relevance.

* advanced knowledge of location characteristics mainly learned from Internet sources.

20

Dataset: Ground Truth #2

Some basic statistics:

- **devset:**

relevance	<i>Kappa agreement*</i> 0.85	<i>% relevant img.</i> 70
diversity	<i>avg. clusters per location</i> 23	<i>avg. img. per cluster</i> 8.9

- **testset:**

relevance	<i>Kappa agreement*</i> 0.75	<i>% relevant img.</i> 67
diversity	<i>avg. clusters per location</i> 23	<i>avg. img. per cluster</i> 8.8

- **credibilityset:**

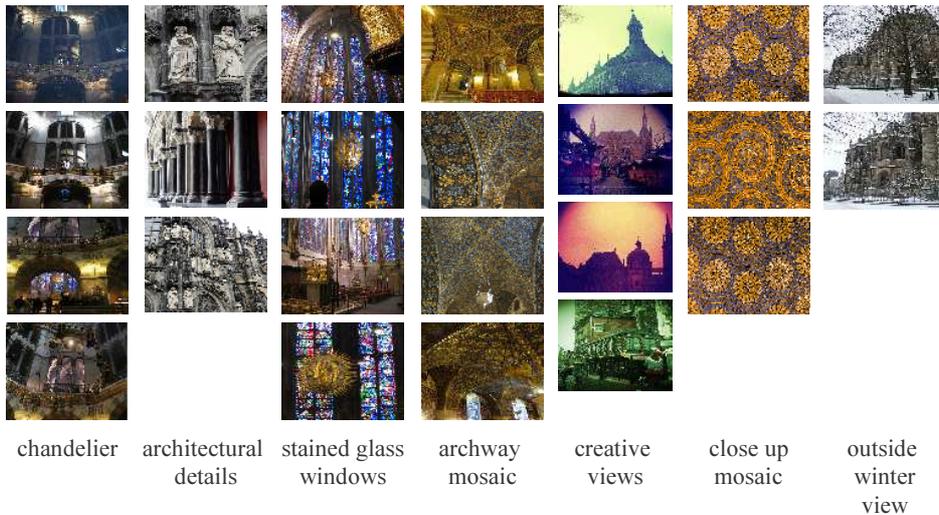
relevance	<i>Kappa agreement*</i> 0.75	<i>% relevant img.</i> 69
-----------	--	------------------------------

*Kappa values > 0.6 are considered adequate and > 0.8 are considered almost perfect.

21

Dataset: Ground Truth #3

Diversity annotation example (Aachen Cathedral, Germany):



22

Evaluation: Required Runs

Participants are required to submit up to 5 runs:

- **required runs:**

- run 1: automated using *visual information only*;
- run 2: automated using *textual information only*;
- run 3: automated using *textual-visual* fused without other resources than provided by the organizers;

- **general runs:**

- run 4: automated using *credibility information*;
- run 5: *everything allowed*, e.g., human-based or hybrid human-machine approaches, including using data from external sources (e.g., Internet).

23

Evaluation: Official Metrics

official ranking F1@20

- **Cluster Recall* @ X = N_c/N (CR@X)**

where X is the cutoff point, N is the total number of clusters for the current location (from ground truth, $N \leq 25$) and N_c is the number of different clusters represented in the X ranked images;

- **Precision @ X = R/X (P@X)**

where R is the number of relevant images;

- **F1-measure @ X = harmonic mean of CR and P (F1@X)**

Metrics are reported for different values of X (5,10,20,30,40 and 50) on per location basis as well as overall (average).

* cluster recall is computed only for the relevant images.

24

Participants: Basic Statistics

- **Survey (February 2014):**
 - 66 (55) respondents were interested in the task, 26 (23) very interested;
- **Registration (April 2014):**
 - 20 (24) teams registered from 15 (18) different countries (3 teams are organizer related);
- **Crossing the finish line (September 2014):**
 - 14 (11) teams finished the task, 12 (8) countries, including 3 organizer related teams (no late submissions);
 - 54 (38) runs were submitted from which 1 (2) **brave human-machine!**
- **Workshop participation (October 2013):**
 - 10 (8) teams are represented at the workshop.

* the numbers in the brackets are from 2013.

25

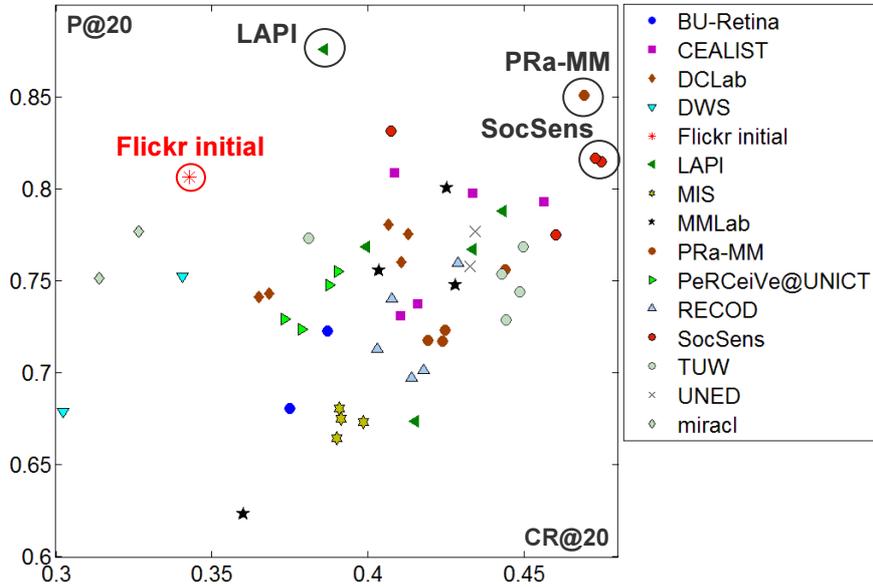
Participants: Submitted Runs

<i>team</i>	<i>country</i>	<i>1-visual</i>	<i>2-text</i>	<i>3-text-visual</i>	<i>4-cred.</i>	<i>5-free</i>
BU-Retina	Turkey	✓	x	x	x	visual
CEALIST*	France, Austria	✓	✓	✓	✓	visual+cred.
DCLab	Hungary	✓	✓	✓	✓	multimodal
DWS	Germany	✓	✓	x	x	x
LAPI*	Romania, Italy	✓	✓	✓	✓	human-mach.
miracl	Tunisia	✓	x	x	x	visual
TUW*	Austria	✓	✓	✓	✓	multimodal
MIS	Austria	✓	✓	✓	x	visual
MMLab	Belgium, S. Korea	✓	✓	✓	x	visual-text
PeRCeiVe@UNICT	Italy	✓	✓	✓	x	visual
PRa-MM	Italy	✓	✓	✓	✓	multimodal
Recod	Brazil	✓	✓	✓	✓	multimodal
SocSens	Greece	✓	✓	✓	x	visual-text
UNED	Spain	x	✓	x	x	text

* organizer related team.

26

Results: P vs. CR @20 (all runs)



27

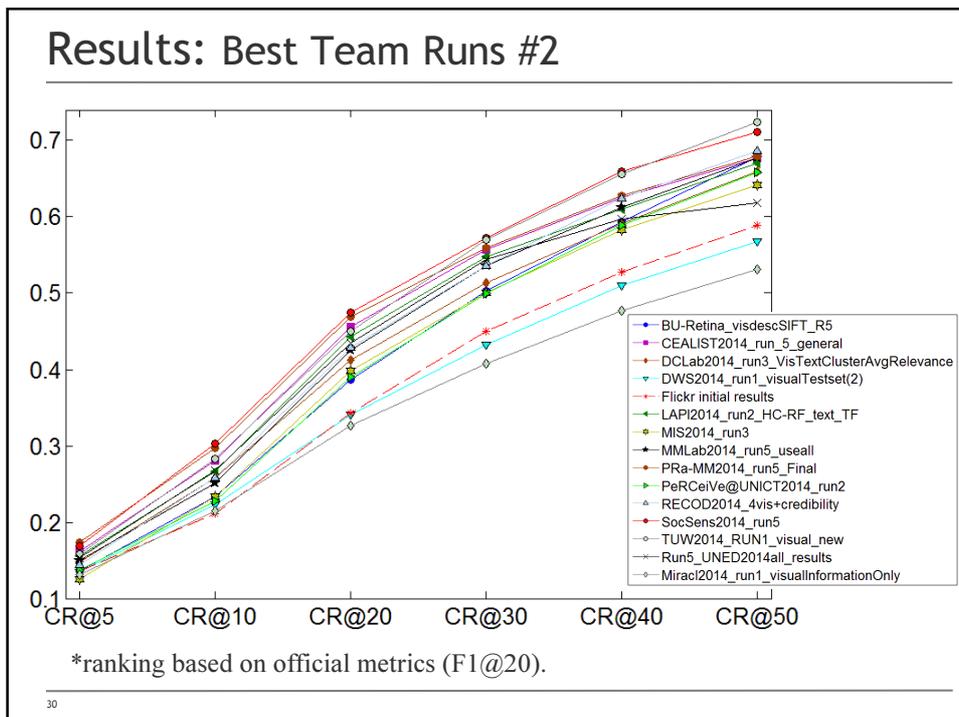
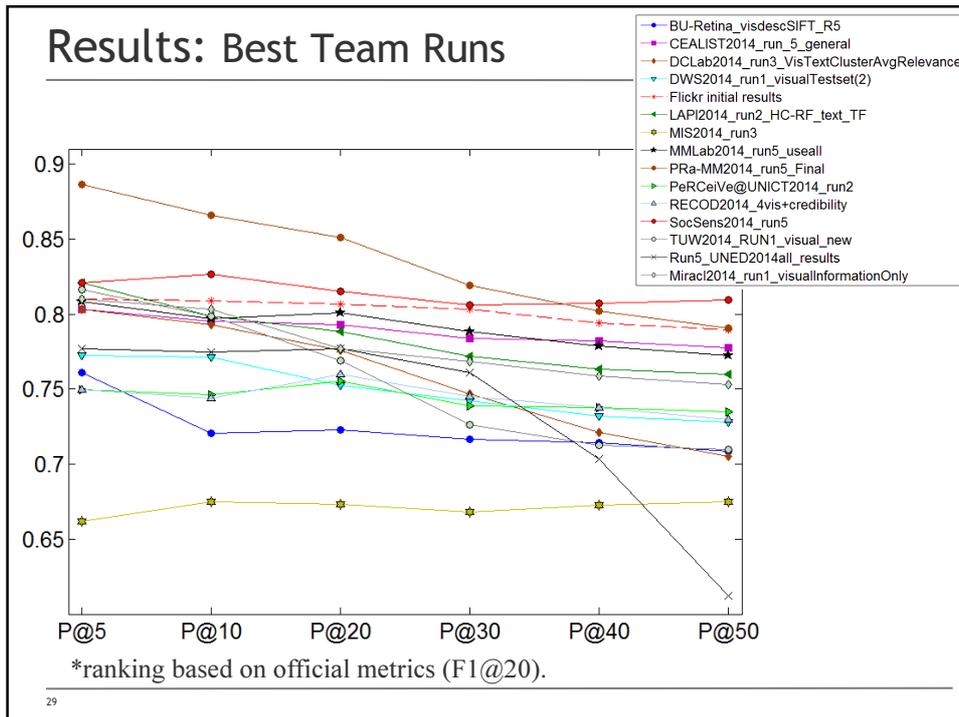
Results: Official Ranking According to F1@20

> team best runs only (full ranking will be sent via email);

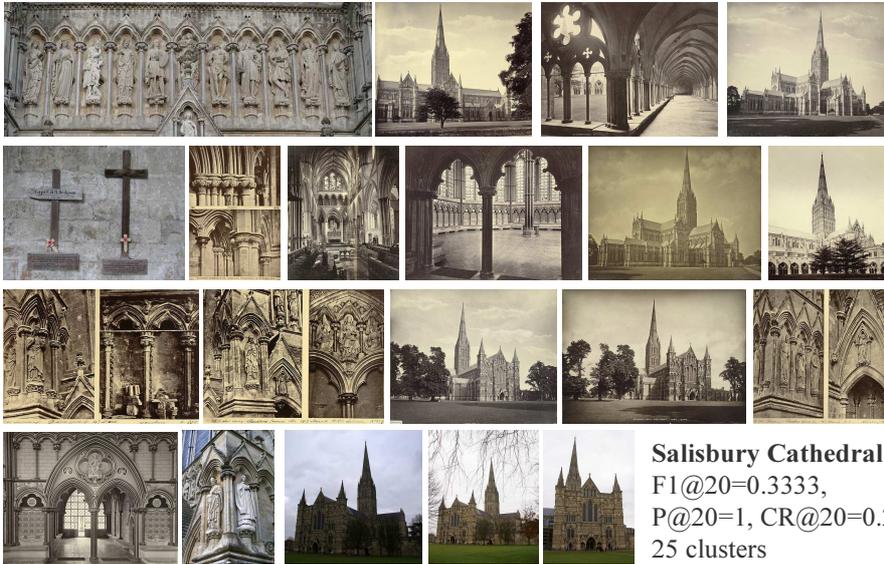
team/run	P@10	P@20	CR@10	CR@20	F1@10	F1@20
PRa-MM2014_run5_Final	0.8659	0.8512	0.2976	0.4692	0.4362	0.5971
SocSens2014_run5	0.8268	0.815	0.3027	0.4747	0.4394	0.5943
CEALIST2014_run_5_general	0.7951	0.7931	0.2803	0.4563	0.4076	0.571
TUW2014_RUN1_visual_new.test	0.7984	0.7687	0.2827	0.4497	0.4124	0.5602
LAPI2014_run2_HC-RF_text_TF	0.7984	0.7882	0.2661	0.4431	0.3928	0.5583
Run5_UNED2014all_results	0.7748	0.7772	0.2679	0.4343	0.3932	0.5502
MMLab2014_run5_useall	0.7967	0.8008	0.2508	0.4252	0.3748	0.5455
RECOD2014_4vis+credibility	0.7439	0.7598	0.2585	0.4288	0.3805	0.5423
DCLab2014_run3_VisTextClusterAvgRelevance	0.7927	0.7756	0.2578	0.4127	0.3838	0.5305
PeRCeiVe@UNICT2014_run2	0.7463	0.7553	0.2271	0.3902	0.3431	0.5063
BU-Retina_visdescSIFT_R5	0.7203	0.7228	0.2339	0.387	0.3492	0.4966
MIS2014_run3	0.6748	0.6732	0.2336	0.3985	0.3433	0.4949
Flickr initial results	0.8089	0.8065	0.2112	0.3427	0.3287	0.4699
DWS2014_run1_visualTestset(2)	0.7715	0.7524	0.2224	0.3405	0.3385	0.46
Miracl2014_run1_visualInformationOnly	0.8033	0.7772	0.2145	0.3265	0.3326	0.4501

Best improvements compared to Flickr (in percentage points): P@20 4.5, CR@20 13.

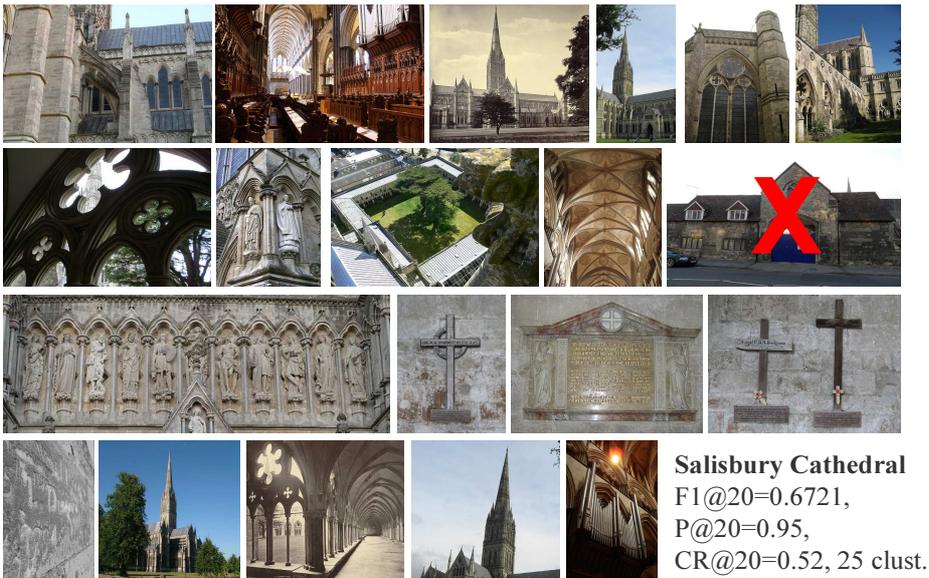
28



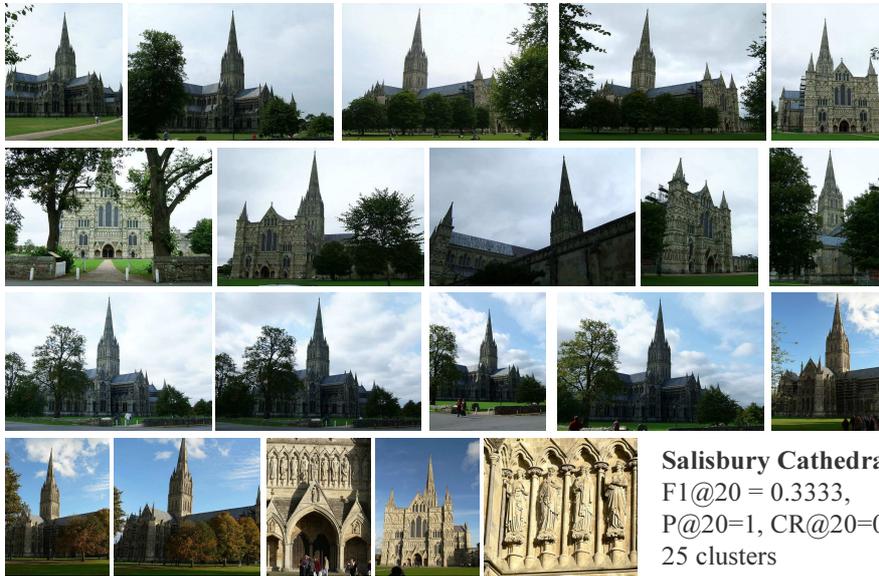
Results: Visual Results - Flickr Initial



Results: Visual Results #2 -Best F1@20



Results: Visual Results #3 - Lowest F1@20



Brief Discussion

Methods:

- this year mainly clustering, re-ranking, optimization-based and relevance feedback (including machine-human);
- best run F1@20: pre-filtering + hierarchical clustering + tree refining + re-ranking using visual-text-cred. information (PRa-MM);
- user tagging credibility information proved its potential and should be further investigated in social retrieval scenarios.

Dataset:

- still low resources for location Creative Commons on Flickr;
- diversity annotation for 300 photos much difficult than for 100;
- descriptors were very well received (employed by most of the participants).

Present & Perspectives

For 2014:

- the task was a full task this year,
- the entire dataset is to be publicly released (soon).



For 2015:

- working on a new use case scenario.



Acknowledgements. Many thanks to:

MUCKE project (Mihai Lupu, Adrian Popescu) for funding the annotation process.

Task auxiliaries: Alexandru Gînscă (CEA LIST, France), Adrian Iftene (Faculty of Computer Science, Alexandru Ioan Cuza University, Romania).

Task supporters: Bogdan Boteanu, Ioan Chera, Ionuț Duță, Andrei Filip, Corina Macovei, Cătălin Mitrea, Ionuț Mironică, Irina Nicolae, Ivan Eggel, Andrei Purică, Mihai Pușcaș, Oana Pleș, Gabriel Petrescu, Anca-Livia Radu, Vlad Ruxandu.

35

Questions & Answers

Thank you!

... and please contribute to the task by
uploading free Creative Commons
photos on social networks! 😊

See you at the poster session and for the
technical retreat ...

36