Listening to the Retriever:
Perturbation-Sensitive Question Selection for Interactive Person Retrieval

Yang Bai[1] · Yunhui Shao[1] · Shuai You[1] · Bin Yang[1] · Jie Qiao[1] · Min Cao[2] · Mang Ye[1]†

[1] School of Computer Science, Wuhan University

[2] School of Computer Science and Technology, Soochow University

NEURAL INFORMATION PROCESSING SYSTEMS 2026 · PARIS

01 Background

- **Text-to-image person retrieval.** Identify a target person in a surveillance gallery using a natural-language description.
- **Task limitation.** The initial description may omit details that distinguish visually similar candidates.
- **Interactive person retrieval**

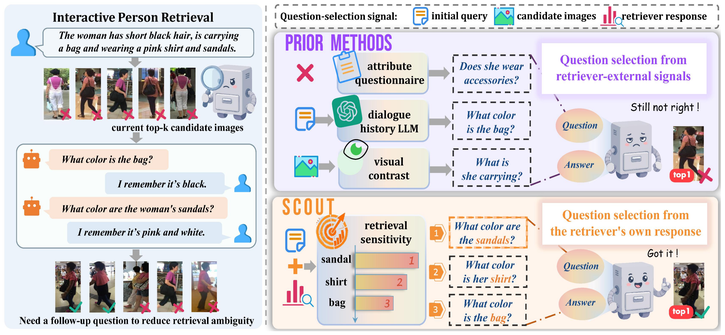
System asks → User answers → Retriever reranks

Which question best distinguishes the current candidates?

02 Key Insight

Generally informative ≠ discriminative for the retriever

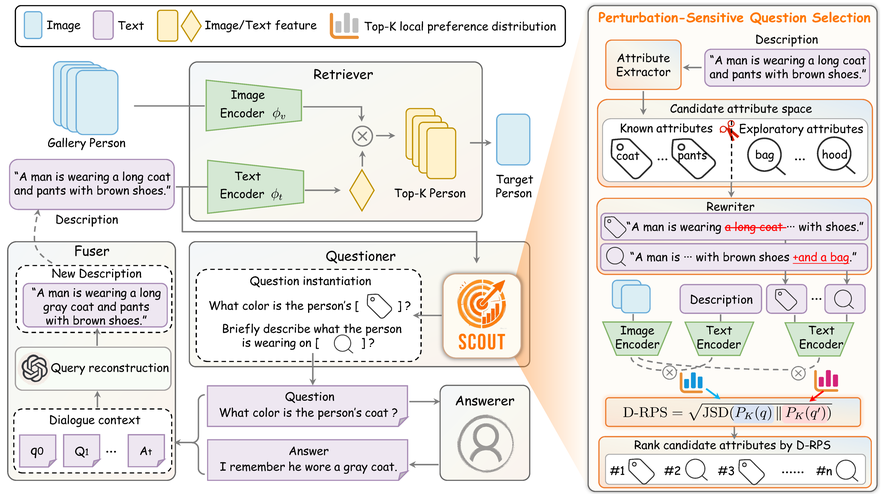
Prior methods use an external signal to select the next question. However, that signal says what is generally informative — **not what is most discriminative for the retriever.**



03 Contributions

- **Signal.** SCOUT shifts question selection from retriever-external signals to the retriever's own response.
- **Strategy.** Retrieval Perturbation Sensitivity (RPS) ranks follow-up questions by how much a semantic perturbation of the query shifts the retrieval preferences.
- **Validation.** RPS-based question selection outperforms existing strategies in ablations. Theoretical analysis connects D-RPS to expected information gain.

04 SCOUT: A training-free interaction loop



05 What to ask

① **Probe.** Given a query q, create a perturbed query q' : delete a *known* attribute or append one for an *exploratory* attribute.

② **Score.** Use D-RPS to quantify the shift in the similarity distribution from q to q'.

$$\text{D-RPS}(q, q') = \sqrt{\text{JSD}\big(P_K(q) \parallel P_K(q')\big)}$$

③ **Ask.** Prioritize cues with high sensitivity scores and turn the selected cue into a follow-up question.

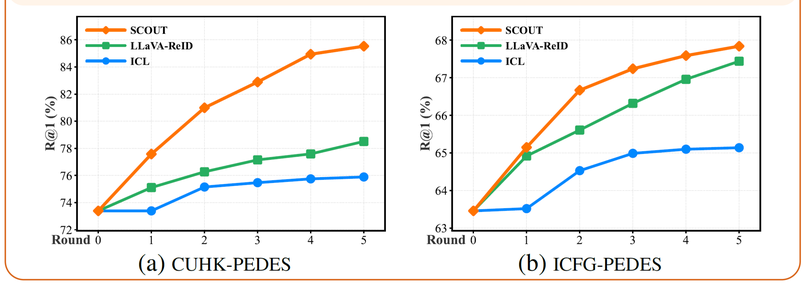
06 Why D-RPS?

- **Bounded change.** $|P_K(q)(A) - P_K(q')(A)| \le \sqrt{2} \cdot \text{D-RPS}(q, q')$. Small D-RPS implies small probability changes for any event on the shared candidate set.
- **Fisher geometry.** $\text{D-RPS}(q, q')^2 = \frac{1}{8}\eta^T(\text{Diag}(p) - pp^T)\eta + o(\|\eta\|^2)$. D-RPS^2 is the Fisher-weighted change of the retrieval distribution along the perturbation direction, reflecting the local sensitivity of the perturbation direction.
- **Information gain.** $\text{JSD}_\pi(\{P_a\}) = I(A; X)$. Under a binary uniform prior, D-RPS reflects the information gain of querying the attribute.

07 Main Results

Methods	Venue	CUHK-PEDES R@1	R@5	R@10	mAP	ICFG-PEDES R@1	R@5	R@10	mAP	RSTPReid R@1	R@5	R@10	mAP
Interactive person retrieval methods													
IRRA[6]	CVPR'23	73.38	89.93	93.71	66.13	63.46	80.25	85.82	38.06	60.20	81.30	88.20	47.17
+ICL[15]	CVPR'25	75.89	89.44	93.62	67.28	65.14	80.11	85.63	38.46	65.35	81.35	88.10	48.13
+LLaVA-ReID[14]	ICML'25	78.51	92.43	95.67	70.61	67.44	82.91	87.69	40.60	69.85	88.10	92.55	54.92
+SCOUT	Ours	85.53	95.14	97.25	75.66	67.84	83.61	88.31	40.54	69.50	87.75	92.35	53.35
RDE[7]	CVPR'24	75.94	90.14	94.12	67.56	67.68	82.47	87.36	40.06	65.35	83.95	89.90	50.88
+ICL[15]	CVPR'25	77.91	90.27	94.14	69.13	69.02	82.45	87.36	41.21	70.55	85.95	91.65	53.68
+LLaVA-ReID[14]	ICML'25	79.39	93.02	95.96	71.70	69.98	84.32	88.60	42.12	72.00	88.95	93.55	56.06
+SCOUT	Ours	86.84	95.39	97.45	76.51	71.36	85.40	89.61	42.43	76.75	90.35	94.20	58.52
HAM[11]	CVPR'25	77.99	91.34	95.03	69.72	69.95	83.88	88.39	42.72	72.50	87.70	91.95	55.47
+ICL[15]	CVPR'25	79.11	91.55	94.74	70.98	70.95	84.28	88.68	44.01	73.90	86.85	91.60	55.99
+LLaVA-ReID[14]	ICML'25	80.95	92.53	95.52	73.34	71.38	84.04	88.43	45.90	75.30	88.20	93.40	60.16
+SCOUT	Ours	**88.94**	**96.43**	**97.87**	**79.11**	**75.46**	**87.74**	**91.49**	**46.39**	**83.75**	**93.30**	**96.00**	**63.20**

08 Gains over interaction rounds



09 Ablations & Visualization

Question Selection Strategy	R@1	R@5	R@10	mAP
baseline	73.38	89.93	93.71	66.13
Random	77.45	91.68	94.95	68.87
Look-forward(LLaVA-ReID)	78.15	91.47	94.72	70.18
Fixed Questionnaire(ICL)	81.25	94.00	96.90	73.06
History-Based LLM(ChatPR)	83.04	94.13	96.77	73.38
Perturbation-Sensitive(Ours)	**85.53**	**95.14**	**97.25**	**75.66**

