



# 《Professional Reading and Writing》 Literature Survey

**Southeast University**

**Professor Shuai Wang**

**School of Computer Science and Engineering,**

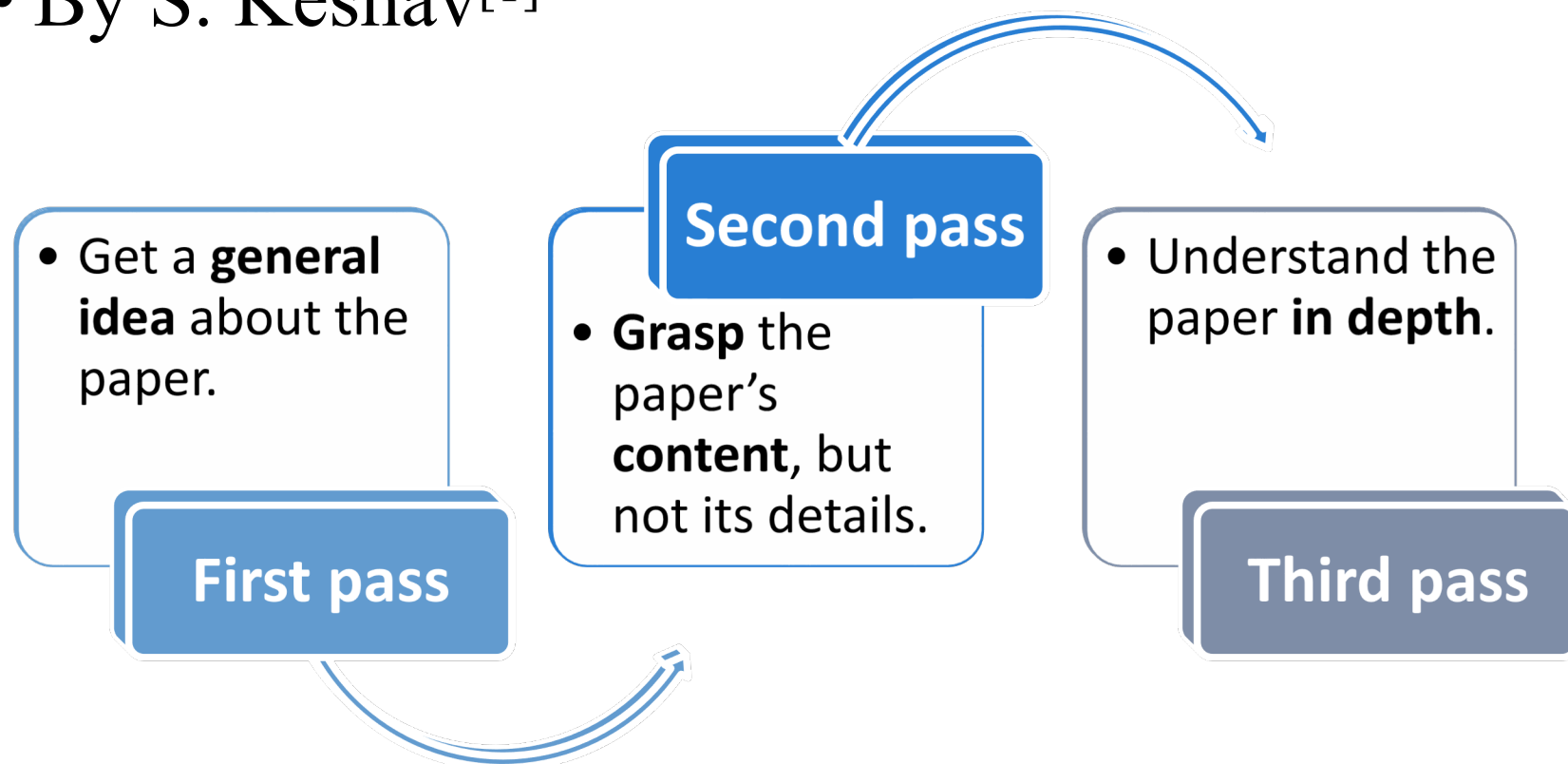
**School of Software Engineering, School of Artificial Intelligence**

**shuaiwang@seu.edu.cn**

**<http://121.40.241.32>**

# Three-Pass Approach to Read a paper

- By S. Keshav<sup>[1]</sup>



[1] S. Keshav, "How to Read a Paper," *ACM SIGCOMM Computer Communication Review* 37(3) : 83-84, July 2007.



# The 1<sup>st</sup> Pass

A **quick scan** to get a bird's-eye view of the paper (5-10 mins).

1. Carefully read the **title**, **abstract**, and **introduction**
2. Read the **section and sub-section *headings***, but ignore everything else
3. Read the **conclusion**
4. Glance over the **references**, mentally ticking off the ones you've already read



# At the End of this Pass ...

1. **Category:** What type of paper is this?
  - A theoretical paper?
  - A system paper ?
  - A survey?
2. **Context:** What is the problem (space)? Which other papers is it related to?
3. **Correctness:** Do the assumptions appear to be valid?
4. **Contributions:** What are the main contributions?
5. **Clarity:** Is the paper well written?



# The 2<sup>nd</sup> Pass

Read with greater care (up to 1 hour)

1. Look carefully at the **figures, diagrams** and other **illustrations** in the paper.
2. **Mark relevant unread references** for further reading
3. Ignore proofs, extensions, and appendix



# At the End of this Pass ...

---

1. Summarize the content of the paper
2. Being able to explain the main ideas of the paper to someone else



# The 3<sup>rd</sup> Pass

Attempt to virtually re-implement the paper (up to 4-5 hours)

1. Make the same assumptions as the authors
2. Re-create the work, re-prove the results, ...
3. think about how you yourself would present a particular idea
4. Compare this re-creation with the actual paper
5. challenge every assumption



# At the End of this Pass ...

---

1. **reconstruct** the entire structure of the paper from memory.
2. identify its **strong and weak points**.
3. Identify hidden failings and assumptions
4. Derive new ideas for future work





# Literature Review of Different Directions in Computer Science



# Different Directions in Computer Science

- 1) Computer architecture/high-performance computing/storage system;
- 2) Computer network;
- 3) Network and cyber security;
- 4) Software engineering/system software/program language
- 5) Database/data mining/information retrieval;
- 6) Theory in computer science;
- 7) Computer graphics and multimedia;
- 8) Artificial intelligence;
- 9) Human-computer interaction and ubiquitous computing;
- 10) Crossing/emerging fields, etc.



# Focusing on a Specified Field

# Doing a Literature Survey

- What is it?
- Required to read tens of papers, perhaps in an unfamiliar field
- Which paper should read?
- 3-pass approach





# Basic methods of field research

- Searching using Keywords
- Reading related work
  - Survey paper
  - Related theory or system papers
- Author and top conference
- Take a look at its Citations



# Keyword

- Keyword is the core of searching paper.
- A good keyword connects a large number of good papers, and even a field.
- Help us quickly identify whether a paper belongs to our field.
- **However, be patient to find keywords, you need to keep reading and accumulating.**
- Keeping on reading papers with different keywords in the same field, your understanding of the field is accumulated.



# Keywords Searching

1. If you know English keywords, find English papers; otherwise using Chinese keywords
2. Read related papers to **expand the keywords**
3. Categorize and refine keywords
  - remove those you are not interested in
4. Communicate with the author, **a good communication can get a lot of the latest keywords**
5. Analyze recent popular papers, update and expand keywords



# References

1. Find a related (important) paper
2. Grasp the basic content by reading it with the three-pass method
3. Focus on “related work” in the paper
4. Read those closely related work, and then their references
5. Think about their relationships and categorize these papers





# Core papers and core researchers

- Find some paper recurring?
- Find some author recurring?
- Congratulations! You found the core papers and researchers in the field



# Author and top conference

- Visit the homepage of the core author
- Find their latest published papers
- which conference the paper was published in
  - This may be the top conference in this field

# Top conference

- Visit the homepages of these top conferences to find the latest collections
- Quickly browse the list of papers and find out your interested papers
- Read the paper (three-pass reading method)
- Continuous the above steps with new related
  - Key words
  - Core paper
  - Key researcher



# Papers are cited and retrieved

- The cited list of papers is another good source of related work
- Using google scholar to find them...



---

# Choosing your interested field

---



# How to choose a research filed

- Your interests
- Enter the lab ahead of time
  - Participate in the group meeting
  - Understand the contents of the laboratory project
- Attend academic report lectures



# Course Report & Presentation



# 总体要求

- 按照各自兴趣查找阅读最新高水平学术论文，总结成文字，并上台给你的同学分享汇报。
  - 分组：3-4名同学一组
  - 阅读：至少5篇论文
  - 写作：5000字左右
  - 汇报：20分钟左右上台报告





# 分组要求

1. 按照兴趣分组
2. 3-4名同学组成一组
3. 确定一个组长



# 课程报告要求

1. 确定小组方向：在CCF推荐中选择一个大类
2. 查找该大类下CCF推荐A类会议列表
3. 从该会议列表中查找近两年的论文列表
4. 从该论文列表中找1篇感兴趣的论文
5. 采用三遍阅读法的第一遍方法读该论文
6. 扩展查找更多相关论文
7. 阅读这些找到的论文，扩展找更多论文
8. 所有论文必须围绕一个相关主题



# 总结领域现状

1. 写一篇5000字以上的总结报告
2. 1个section概述（科普）你选择的主题
3. 梳理同一主题论文之间（至少5篇）的关系（可用图或表）
4. 1个section包括5个subsection描述5篇论文
5. 思考这些工作的意义，未来还有什么可以做的

第4和5点要重视，请认真梳理和思考



# 课程报告要求

1. 使用Latex工具写课程报告
2. 使用IEEEtran模板的conference选项
3. 使用中文
  - 编译器选XeLaTeX
  - tex文件中添加`\usepackage[UTF8]{ctex}`
4. 至少1张图
5. 至少1个表格
6. 至少1个带编号公式
7. 至少5篇参考文献
8. 表明小组成员姓名、学号等信息
9. 提交一个PDF文件



# 汇报内容要求

1. 做一个学术报告PPT
2. 以你自己身边的同学作为报告听众  
– 你的同学听得懂
3. 包括你选择的主题（或小方向）的科普性介绍
4. 梳理同一主题论文（5篇左右）之间的关系，  
用图或表的方式展现
5. 重点介绍其中1~2篇论文（不要多于2篇）
6. 包括这些工作的意义或者你觉得未来还可以开展的工作



# 汇报其他要求

1. 图文并茂**切忌大段文字**
  - 避免大量公式和证明
2. 报告时间20分钟左右（包括2-3分钟提问）
3. 要求每小组派代表报告
4. 按照学术报告的风格制作PPT
  - 不要太花哨，专注讲清楚内容
5. **7月26日（含）之前提交PPT电子版**
6. **8月1日（含）之前提交报告**



# Discussion about homework 2

- 根据CCF推荐列表中计算机各个子方向
  - 选择一个你喜欢的子方向
  - 对该子方向进行科普性调研
- 确定子方向后，选择你认为最合适的
  1. 一篇近两年有意思的高水平论文
  2. 一个相关高水平会议
  3. 一个相关高水平期刊
  4. 一位相关领域专家
- 并给出选择的理由



# Discussion about homework 2

- 整理文档，提交如下内容
  1. 选择的子方向名称及对该方向的认知总结
  2. 论文信息（以参考文献格式）
  3. 会议名称和网站链接
  4. 期刊名称和网站链接
  5. 专家名字和主页链接
  6. 上述5个选择的简要理由



# Some examples

## A. 选择方向

我选择的子方向是计算机图形学与多媒体。

认知总结：计算机图形学（英语：computer graphics，缩写为 CG）是研究计算机在硬件和软件的帮助下创建计算机图形的科学学科，是计算机科学的一个分支领域，主要关注数字合成与操作视觉的图形内容。虽然这个词通常被认为是指三维图形，事实上同时包括了二维图形以及影像处理。多媒体（Multimedia），在电脑应用系统中，组合两种或两种以上媒体的一种人机交互式资讯交流和传播媒体。使用的媒体包括文字、图片、照片、声音（包含音乐、语音旁白、特殊音效）、动画和影片，以及程序所提供的互动功能。即计算机图形学与多媒体主要关注的是如何用计算机解决图像、文本、声音的问题。

理由：计算机图形学与多媒体目前应用广泛，在游戏行业、音视频行业等都有很大的应用。并且随着 5G 时代的来临，网络传播的速度越来越快、越来越稳定，短视频行业与自动驾驶行业开始蓬勃发展。而这两个行业都与该领域息息相关，在未来，如何快速、高效处理以不同载体为媒介的信息将至关重要。而我本身就对多媒体有关的领域十分感兴趣，所以此次作业选择深入理解这个方向。



# Some examples

## B. 高水平论文

我选择的文章是 [1]Y. R. Choi and R. M. Kil, "Face Video Retrieval Based on the Deep CNN With RBF Loss," in IEEE Transactions on Image Processing, vol. 30, pp. 1015-1029, 2021, doi: 10.1109/TIP.2020.3040847.

理由: 目前人脸识别已经在世界有了广泛的应用。生活中我们时常能见到在利用人脸识别技术验证你的身份, 只有当我们验证通过我们才有可能能够通过一些例如动车站或者时其他站点的关卡。除此之外, 人脸识别技术还在追击嫌疑犯上发挥了巨大的作用, 而中国准确度最高的人脸识别技术被应用于天眼系统。

这篇文章提出了一个新的框架来提取人脸视频的高度紧凑和判别特征使用深度卷积神经网络的检索任务(美国有线电视新闻网)。人脸视频检索任务是找到视频包含来自数据库的特定人脸与查询相同的人的面部图像或面部视频。关键挑战是用小存储提取判别特征来自具有较大类内变化的面部视频的空间通过不同的角度、光照和面部表情。在最近多年来, 基于 CNN 的二进制哈希和度量学习方法在图像/视频检索任务方面取得了显着进展。然而, 现有的基于 CNN 的二进制哈希和度量学习在不可避免的信息方面有局限性分别是损失和存储效率低下。为了应对这些问题, 提出的框架由两部分组成: 第一, 使用径向基函数核 (RBF) 的新型损失函数损失) 被引入训练神经网络以





# Some examples

## C. 高水平会议

名称: IEEE Visualization Conference

网址: <http://dblp.uni-trier.de/db/conf/visualization/index.html>

理由: IEEE Visualization Conference (VIS) 是计算机图形学与多媒体的一个重要会议。根据 2016 年 Google Scholar 的 h-index 指标排名, VIS 是视觉研究会议评分最高的会议, 也是计算机图形学评分第二高的会议。它在澳大利亚 ICT 会议排名中被评为“A”级, 以及巴西教育部的“A”评级。该会议具有高度选择性, 所有论文的接受率通常小于 25%。这是一场在该领域认可度极高的会议, 所以我选择这场会议。



# Some examples

## D. 高水平期刊

名称: IEEE Transactions on Image Processing

网址: <http://dblp.uni-trier.de/db/journals/tip/>

理由: IEEE Transactions on Image Processing 主要发表关于图像处理和计算机视觉领域的学术论文。IEEE Transactions on Image Processing 的发布是经过经验丰富的科学同行严格审查后的结果, 所以具有其发布的论文具有非常高的价值, 且在业界认可度较高。该杂志鼓励各地研究所投稿, 他们主要关注投稿论文中的创新性以及实际意义。所以, 这是一份十分有价值的期刊。

# Some examples

## E. 相关领域专家

名字: 李飞飞

主页地址: <https://profiles.stanford.edu/fei-fei-li>

理由: 李飞飞博士是斯坦福大学计算机科学系的首任红杉教授。她于 2013 年至 2018 年担任斯坦福大学 AI 实验室主任。2017 年 1 月至 2018 年 9 月在斯坦福大学休假期间, 她担任谷歌副总裁, 并担任谷歌云 AI/ML 首席科学家。李飞飞博士于 1999 年获得普林斯顿大学物理学学士学位, 2005 年获得加州理工学院 (Caltech) 电气工程博士学位。2009 年加入斯坦福大学担任助理教授。在此之前, 她曾在普林斯顿大学 (2007-2009) 和伊利诺伊大学香槟分校 (2005-2006) 任教。

李飞飞博士目前的研究兴趣包括受认知启发的人工智能、机器学习、深度学习、计算机视觉和人工智能 + 医疗保健, 尤其是用于医疗保健服务的环境智能系统。过去, 她还致力于认知和计算神经科学。李博士在 Nature、PNAS、





# Some examples

## REFERENCES

- [1] Young Rok Choi and Rhee Man Kil. “Face Video Retrieval Based on the Deep CNN With RBF Loss”. In: *IEEE Transactions on Image Processing* 30 (2020), pp. 1015–1029.



# Some examples

## 1. 网络与信息安全

网络信息安全是一门涉及计算机科学、网络技术、通信技术、密码技术、信息安全技术、应用数学、数论、信息论等多种学科的综合性学科。它主要是指网络系统的硬件、软件及其系统中的数据受到保护，不受偶然的或者恶意的原因而遭到破坏、更改、泄露，系统连续可靠正常地运行，网络服务不中断。

## 2. 论文

[1]唐迪, 顾健, 俞优, 杨元原. 基于等级保护的个人信息安全分级方法研究[A], 公安部第三研究所. 2020 年“网络安全技术与应用创新”研讨会论文集[C], 公安部第三研究所:《信息网络安全》北京编辑部, 2020:4.

## 3. 会议

2020 年“网络安全技术与应用创新”研讨会

<https://navi.cnki.net/KNavi/DPaperDetail?pcode=CIPD&lwjcode=XXWL202011001&hycode=028812>

## 4. 期刊

《网络安全技术与应用》

<https://wlaq.cbpt.cnki.net/>

## 5. 专家名称

吴春明

<https://person.zju.edu.cn/0095168>

## 6 理由

网络与信息安全与我所学的专业网络空间安全密不可分，而网络与信息安全具体到个人



# Some examples

对数据的处理和利用自古有之，在计算机出现后，大规模管理数据已经成为可能，数据库的概念便自然地出现了，数据库相当于一个虚拟的文件柜，其中的数据按一定的数据模型组织、描述和存储，具有较小的冗余度、较高的数据独立性和易扩展性，并可为各种用户共享，我们可以借此进行海量数据的管理与使用，数据库的使用也是我们的研究方向。当代的大数据，云计算技术等都离不开数据库的利用。我们要做的，利用数据库，增强数据的复用性和联系性，数据的泄露也要列入考虑范围内（虽然这已经无法避免），数据管理技术要与多想技术结合，拓宽信息共享的平台，发现数据利用的更优解。

论文：【1】柴成亮 李国良 赵天宇 骆昱宇 于明鹤 众包数据库综述【J】计算机学报 2020 43 (05)

首先介绍众包的概念;然后介绍设计众包数据库时需考虑的一些基本技术;接着介绍几种主流的众包数据库系统,并介绍该领域未来的研究方向与挑战,内容比较全面。研究内容具有现实性和可操作性。经过对论文的审核可以看出,作者在资料和案例收集上花了不少功夫。得到国家自然科学基金; 国家“九七三”重点基础研究发展计划项目基金资助, 也增加了论文的可靠性。

第十二届中国卫星导航年会 <https://www.beidou.org/> 在今年有着多篇对于数据库应用的论文, 研究了数据库在卫星定位方面如定位系统, 北斗的应用。

计算机学报 <http://cjic.ict.ac.cn/> 是中国计算机领域权威性学术刊物。其宗旨是报道中国计算机领域研究成果, 虽然是一种包含计算机各个领域的期刊, 但是关于数